# THEORY OF (NON-LINEAR) STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS AND ITS APPLICATIONS TO INTEREST RATES

**by**

**TORSTEIN NILSSEN**

***THESIS***

*for the degree of*

***MASTER IN MATHEMATICS***

*(Master of Science)*

*Faculty of Mathematics and Natural Sciences*
*University of Oslo*

*December 2009*

# Preface

Stochastic partial differential equations (SPDEs) have been studied since the 1960s and as Michael Röckner puts it, non-linear SPDEs can be used to model "All kinds of dynamics with stochastic influence ...". The setup is to regard a SPDE as an infinite dimensional valued stochastic differential equation and this thesis presents two approaches to analysing solutions; the variational approach and the semi-group approach.

The content is based on [PR07] plus the notes from a course on SPDEs held by Tusheng Zang at University of Oslo in Spring 2007 (notes taken by An Ta Thi Kieu). For Section 3.2 and the final chapter, I have used notes from a course on interest rates and SPDEs held by Frank Proske at University of Oslo in Spring 2009.

The first chapter deals with integration, differentiation and stochastic integration in infinite dimensions. My work here has been to transfer basic results on the Bochner integral into the Pettis integral. Also I have proved existence of conditional expectation using a generalized form of the Radon-Nikòym theorem to make it more compatible with the Pettis integral. Stochastic integration is simplified to the case of cylindrical Brownian motion.

The second Section introduces some theory from PDEs. The result on Gelfand triples is done by me. Definitions of weak derivatives and Sobolev spaces is included to make the thesis more self contained. The theorem and proof on deterministic equations is based on notes from the course held by Tusheng Zang, but put in a less general setting (which fits better in what follows).

The third chapter is the core of the thesis as it deals with the mentioned infinite dimensional equations of stochastic type. The proof of the Itô formula is a sketch of the proof in [PR07]. In Section 3.2, on mild solutions, I have taken notes from the course held by Frank Proske and generalized the proof from $p = 2$ into $p \geq 2$. Section 3.3 generalizes the result from 2.4 to a result on linear SPDEs. The work here is based on the notes from the course held by Tusheng Zang. The non-linear result in Section 3.4 is taken from [PR07] and is presented here as a sketch. Frank Proske gave me the idea of generalizing the theorem in [BØP05], and so, in Section 3.5 I have proved an existence and uniqueness result on backward SPDEs which includes a class of semi-linear differential operators.

The final chapter is a short chapter on the connection between SPDEs and interest rates. Here I have presented two finite-dimensional models for interest rates, and one infinite-dimensional model. The results in this chapter comes from the course on interest rates by Frank Proske and from [CT06], but is presented here with proofs not found in [CT06].

### Acknowledgement

First and foremost, I would like to thank my supervisor Frank Proske for his enthusiasm and invaluable wisdom; your help has been truly appreciated.

Also, I would like to thank; everybody at B606 and B601 for their wonderful distractions and all the fun we've had; John Christian Ottem, Nikolay Qviller, Elin Røse and Ketil Tveiten for proofreading; my brother Trygve for proof reading, fruitful discussions (thanks to his wife, Siri, for patience) and all the motivation and support I've received for as long as I can remember; An Ta Thi Kieu for the notes from Tusheng Zangs course; Giulia Di Nunno and Nadia Larsen for highly appreciated lectures in stochastic analysis and functional analysis respectively.

Finally, thanks to my wonderful girlfriend Ellen for her constant love and support.

# Contents

# 1 Calculus for Vector-valued Functions

This section deals with integration and differentiation of functions with values in a vector space, or more specifically a Banach space.
For a finite-dimensional vector space, and a function

$$f : \mathcal{K} \to \mathbb{R}^n$$

defined on a set $\mathcal{K}$, one could consider $(f_1, \ldots, f_n)$ as a vector of one-dimensional functions and build an integration theory around this. For an infinite-dimensional Banach space, one has to use the continuous linear functionals defined on this space, as these correspond to the one-dimensional projections on the space.

## 1.1 Pettis Integral

**Definition 1.1.** *Let $(\mathcal{K}, \mathcal{C}, \mu)$ be a finite measure space, and $V$ a real separable Banach space. A function $f : \mathcal{K} \to V$ is called measurable if the composition*

$$\varphi \circ f : \mathcal{K} \to \mathbb{R}$$

*is $\mathcal{C}$-measurable, for all $\varphi \in V^*$.*

**Definition 1.2.** *Let $f : \mathcal{K} \to V$ be a measurable function. If there exists a vector $z \in V$ such that for any $\varphi \in V^*$,*

$$\langle \varphi, z \rangle = \int \langle \varphi, f \rangle d\mu.$$

*The vector $z$ is called the Pettis integral of $f$ and is denoted by $\int f d\mu$.*

**Theorem 1.3.** *If the function $\|f(\cdot)\| : \mathcal{K} \to \mathbb{R}$ belongs to $L^1(\mathcal{K}, \mathcal{C}, \mu)$, then there exists a unique Pettis integral of $f$ which satisfies*

$$\| \int f d\mu \| \leq \int \|f\| d\mu. \tag{1}$$

*Proof.* Define a functional on $V^*$ by

$$\begin{array}{rccc} T: & V^* & \longrightarrow & \mathbb{R} \\ & \varphi & \mapsto & \int \langle \varphi, f \rangle d\mu. \end{array}$$

Since $|\int \langle \varphi, f \rangle d\mu| \leq \int |\langle \varphi, f \rangle| d\mu \leq \|\varphi\| \int \|f\| d\mu$ which is finite by hypothesis, $T$ is a well-defined functional on $V^*$. Look now at $V^*$ with the $w^*$-topology. Since $V$ is assumed to be separable, this topology is induced by the metric

$$d(\psi, \varphi) = \sum_{n=1}^{\infty} |\langle \psi - \varphi, x_n \rangle| 2^{-n}$$

where $\{x_n\}$ is dense in $V$. It follows that the topology is uniquely determined by sequences. Let $\{\varphi_n\}$ be a sequence in $V^*$ converging in the $w^*$-topology to $\varphi \in V^*$. By the Banach-Steinhaus theorem, $\sup_n \|\varphi_n\| < \infty$. The sequence $\langle \varphi_n, f\rangle$ converges almost everywhere to $\langle \varphi, f\rangle$ and since the sequence is dominated by $\sup_n \|\varphi_n\|\|f\|$ which is integrable, it follows that

$$\lim_{n\to\infty} \int \langle \varphi_n, f\rangle d\mu = \int \langle \varphi, f\rangle d\mu$$

so that $T$ is continuous in the $w^*$-topology. Then there exists a $z \in V$ such that $\langle T, \varphi\rangle = \langle \varphi, z\rangle$ which is the desired vector.

Since $V^*$ separates points in $V$, the operation $f \mapsto \int f d\mu$ is well-defined.

Finally, to see (1), by the Hahn-Banach extension theorem, choose $\varphi \in V^*$ such that

$$\|\int f d\mu\| = \langle \varphi, \int f d\mu\rangle = \int \langle \varphi, f\rangle d\mu \leq \int \|f\| d\mu.$$

□

Let $L^1(\mathcal{K}, \mathcal{C}, \mu; V)$ denote the space of Pettis-integrable functions with values in $V$. When no confusion can arise, the space will be denoted $L^1(\mathcal{K}; V)$.

**Example 1.4.** *Let $V = \mathbb{R}^n$, for some $n \in \mathbb{N}$. Since $(\mathbb{R}^n)^* = span\{\pi_j : j = 1, \dots n\}$, where $\pi_j : \mathbb{R}^n \to \mathbb{R}$ denotes the projection onto the $j$-th coordinate, a function $X : \Omega \to \mathbb{R}^n$ is a random variable if and only if all of its coordinates, $X^j$, are (standard) random variables. Also, the expectation is a vector, given by $E[X] = (E[X^1], \dots, E[X^n])$.*

**Example 1.5.** *Let $f \in L^1(\mathcal{K}; \mathcal{H})$, for a separable Hilbert-space $\mathcal{H}$ with orthonormal basis $\{e_n\}$. Then the integral has the representation*

$$\int f d\mu = \sum_{n=1}^{\infty} \left( \int \langle f, e_n\rangle d\mu \right) e_n.$$

*Proof.* Identifying $\mathcal{H}^*$ with $\mathcal{H}$ via the Riesz identification map $y \mapsto \langle \cdot, y\rangle$ and using that for $x \in \mathcal{H}$ it holds $x = \sum_{n=1}^{\infty} \langle x, e_n\rangle e_n$, it follows

$$\int f d\mu = \sum_{n=1}^{\infty} \langle \int f d\mu, e_n\rangle e_n = \sum_{n=1}^{\infty} \left( \int \langle f, e_n\rangle d\mu \right) e_n.$$

□

The latter example shows that, as expected from Example 1.4, the infinite-dimensional integral can be considered as an infinite sequence of one-dimensional integrals.

Similarly one defines the extension of $L^p$ spaces as

$$L^p(\mathcal{K}; V) = \{f : \mathcal{K} \to V \mid f \text{ is measurable and } \|f\| \in L^p(\mathcal{K})\}.$$

**Theorem 1.6.** *The space $L^p(\mathcal{K};V)$ is a Banach space.*

The proof is based on the usual Riesz-Fischer theorem from [Bar95] (where the one-dimensional case is considered).

*Proof.* Let $\{f_n\} \subset L^p(\mathcal{K};V)$ be a Cauchy sequence and choose a subsequence (still indexed by $n$) such that $\|f_{n+1} - f_n\|_p \leq 2^{-n}$. Define

$$g(\omega) = \|f_1(\omega)\| + \sum_{n=1}^{\infty} \|f_{n+1}(\omega) - f_n(\omega)\|.$$

Then by Fatou's lemma

$$\left(\int g^p d\mu\right)^{1/p} \leq \liminf_{k\to\infty} \left(\|f_1\|_p + \sum_{n=1}^{k} \|f_{n+1} - f_n\|_p\right) \leq \|f_1\|_p + 1,$$

so that $g \in L^p(\mathcal{K})$. Then let $F = \{g < \infty\}$, which has full measure and define the $V$-valued function

$$f(\omega) = \begin{cases} f_1(\omega) + \sum_{n=1}^{\infty} f_{n+1}(\omega) - f_n(\omega) & \text{if } \omega \in F \\ 0 & \text{otherwise .} \end{cases}$$

Since $V$ is a Banach space the limit exists and $\{f_n\}$ converges $\mu$-a.s. everywhere to $f$. Since $\|f_n\| \leq g$ it follows by the dominated convergence theorem that $f \in L^p(\mathcal{K};V)$ and

$$\int \|f - f_n\|^p d\mu \to 0$$

as $n \to \infty$. $\square$

This shows that $L^p(\mathcal{K};V)$ is the perfect generalization of the one-dimensional case. A natural question could now be if the celebrated Radon-Nikodŷm theorem holds. In most cases the answer is positive, but let us first introduce some terminology that will be useful:

**Definition 1.7.** *Let $\mathcal{C}$ be a $\sigma$-algebra of sets of $\mathcal{K}$. A set function $\nu : \mathcal{C} \to V$ where $V$ is a Banach space, is called a vector-measure if, for any disjoint sequence of sets $\{F_j\}$, it holds that*

$$\nu\left(\bigcup_{j=1}^{\infty} F_j\right) = \sum_{j=1}^{\infty} \nu(F_j),$$

*where the right hand side converges in the norm topology.*
*When $\nu$ satisfies*

$$\|\nu\| := \sup_{\{F_j\}_{j=1}^m \in \mathcal{D}} \sum_{j=1}^{m} \|\nu(F_j)\|_V < \infty$$

*where $\mathcal{D}$ is the family of all finite partitions of $\mathcal{K}$, the vector measure $\nu$ is said to be of finite variation.*

**Definition 1.8.** *Let $(\mathcal{K}, \mathcal{C}, \mu)$ be a finite measure space. A Banach space $V$ is said to have the Radon-Nikodỳm property with respect to $\mu$ if for every vector-measure $\nu : \mathcal{C} \to V$ with bounded variation such that*

$$\mu(F) = 0 \quad \Rightarrow \quad \nu(F) = 0 \text{ (the zero-vector)}$$

*there exists a $g \in L^1(\mathcal{K}; V)$ such that*

$$\nu(F) = \int_F g \, d\mu.$$

There exists separable Banach spaces and vector measures such that the Radon-Nikodỳm property does not hold. Fortunately, the following theorem provides a sufficient result for the Banach spaces that will be used. The proof can be found in [DU77]

**Theorem 1.9.** *Every reflexive Banach space has the Radon-Nikodỳm property for any vector measure.*

It is well known that $L^p(\mathcal{K})^* = L^q(\mathcal{K})$ (where $\frac{1}{p} + \frac{1}{q} = 1$) in the one-dimensional case and a further question can be if this holds more generally. One inclusion is easily shown. Namely, let $g \in L^q(\mathcal{K}; V^*)$ and define $\varphi_g$ on $L^p(\mathcal{K}; V)$ by

$$\langle \varphi_g, f \rangle = \int \langle g, f \rangle d\mu. \tag{2}$$

By Hölder's inequality $|\langle \varphi_g, f \rangle| \leq \|g\|_q \|f\|_p$, so that $\varphi_g$ is a continuous linear functional on $L^p(\mathcal{K}; V)$, and $\|\varphi_g\| \leq \|g\|_q$. In fact, the following result is proved in [DU77]:

**Lemma 1.10.** *Define $\varphi_g \in (L^p(\mathcal{K}; V))^*$ as in (2). Then*

$$\|\varphi_g\| = \|g\|_p. \tag{3}$$

This shows that $g \mapsto \varphi_g$ is an isometry of $L^p(\mathcal{K}; V)$ into $(L^q(\mathcal{K}; V))^*$. For spaces with the Radon-Nikodỳm-property the following hold.

**Theorem 1.11.** *Assume that $V$ is reflexive. Then*

$$(L^p(\mathcal{K}; V))^* = L^q(\mathcal{K}; V^*).$$

To prove this, the following result is needed which can be found in [PZ92].

**Lemma 1.12.** *Let $f : \mathcal{K} \to V$ be a measurable function. There exists a sequence of step functions $\{f_n\}$, i.e.*

$$f_n = \sum_{j=1}^{m_n} v_j \chi_{F_j}$$

*for sequences $\{v_j\} \subset V$ and $\{F_j\} \subset \mathcal{C}$ such that the sequence $\|f_n(\omega) - f(\omega)\|$ is monotonically decreasing for every $\omega \in \mathcal{K}$ .*

Using this lemma and dominated convergence, it follows that for $f \in L^p(\mathcal{K}; V)$ there exists a sequence of step functions, $\{f_n\}$ (which lies in $L^p(\mathcal{K}; V)$ since $\mu(\mathcal{K}) < \infty$), such that

$$\int \|f_n - f\|^p d\mu \to 0$$

as $n \to \infty$.

*Proof of 1.11.* Let $\varphi \in (L^q(\mathcal{K}; V))^*$ and define the map

$$\psi : \mathcal{C} \times V \to \mathbb{R}$$

by $\psi(F, v) = \varphi(\chi_F v)$. Then, for fixed $F \in \mathcal{C}$, $\psi(F, \cdot)$ is a linear map on $V$. Also, for a $v$ in the unit ball of $V$,

$$|\psi(F, v)| \leq \|\varphi\||\mu(F)|^{1/p}$$

so that $\psi(F, \cdot) \in V^*$. Then the map $F \mapsto \psi(F, \cdot)$ is a $V^*$-valued vector measure, by the continuity of $\varphi$. To see that $F \mapsto \psi(F, \cdot)$ is of bounded variation, let let $\epsilon > 0$ and $\{F_1, \ldots, F_n\}$ be a partition of $\mathcal{K}$. Choose $\{v_1, \ldots v_n\}$ in the unit ball of $V$ such that

$$\|\psi(F_k, \cdot)\| \leq \psi(F_k, v_k) + \frac{\epsilon}{n}.$$

Then

$$\sum_{k=1}^{n} \|\psi(F_k, \cdot)\| \leq \sum_{k=1}^{n} \psi(F_k, v_k) + \epsilon \leq \varphi(\sum_{k=1}^{n} \chi_{F_k} v_k) + \epsilon \leq \|\varphi\|\mu(\mathcal{K})^{1/p} + \epsilon$$

so that

$$\|\psi(\cdot, \cdot)\| \leq \|\varphi\|\mu(\mathcal{K})^{1/p} + \epsilon$$

and hence $\|\psi(\cdot, \cdot)\| \leq \|\varphi\|\mu(\mathcal{K})^{1/p}$ since $\epsilon$ was arbitrary. As $V$ has the Radon-Nikodŷm-property there exists a $g \in L^1(\mathcal{K}; V^*)$, such that

$$\varphi(\chi_F v) = \int_F \langle g, v \rangle d\mu. \tag{4}$$

Let $F_k = \{\|g\|_{V^*} \leq k\}$ and define the localization of $g$ by $g_k := g\chi_{F_k}$. Since $\mu(\mathcal{K}) < \infty$, $g_k \in L^q(\mathcal{K}; V^*)$. Define the restriction $\varphi_k := \varphi|_{L^p(F_k; V)}$. Then $\|\varphi_k\| \leq \|\varphi\|$ and by linearity of (4) it holds that

$$\varphi_k(f) = \int \langle g_k, f \rangle d\mu \tag{5}$$

for all step functions. Let $f \in L^p(F_k; V)$ be arbitrary. Choose a sequence of functions as in Lemma 1.12. Then, since $\varphi_k$ is continuous, $\varphi_k(f_n) \to \varphi_k(f)$, and by Hölder's inequality

$$\int |\langle g_k, f - f_n \rangle| d\mu \leq \|g_k\|_q \|f - f_n\|_p \to 0$$

so that (5) extends to $L^p(F_k;V)$. Then by (3), $\|g_k\|_q = \|\varphi_k\| \leq \|\varphi\|$, and by Fatou's Lemma

$$\int \|g\|_{V^*}^q d\mu \leq \liminf_{k\to\infty} \|\varphi_k\|^q \leq \|\varphi\|^q$$

which shows that $g \in L^q(\mathcal{K};V^*)$ and arguing similarly as above,

$$\varphi(f) = \int \langle g, f\rangle d\mu$$

for all $f \in L^p(\mathcal{K};V)$.

$\square$

In the proof, the idea of using localization of $g$ by $g_k$ is taken from [DU77].

It is also possible to prove the theorem by use of tensor products. This can be done by identifying $L^p(\mathcal{K};V)$ with $L^p(\mathcal{K}) \otimes V$ using Lemma 1.12. Now $(X \otimes Y)^* \simeq X^* \hat{\otimes} Y^*$ for the right choice of topologies, and the result follows.

The above proof is a more measure theoretic proof, and generalizes the one-dimensional case perfectly.

## 1.2 Conditional Expectation

**Theorem 1.13.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{G} \subset \mathcal{F}$ be a sub-$\sigma$-algebra. Let $X \in L^1(\Omega, \mathcal{F}, P; V)$. Then there exists a $P$-a.s. unique $\mathcal{G}$-measurable function*

$$E[X|\mathcal{G}] : \Omega \to V$$

*such that*

$$\int_G E[X|\mathcal{G}]dP = \int_G XdP$$

*for all $G \in \mathcal{G}$. Also it holds that*

$$\|E[X|\mathcal{G}]\| \leq E[\|X\||\mathcal{G}], \quad P-a.s. \tag{6}$$

*Proof.* Let $\nu : \mathcal{G} \to V$ be defined by $\nu(G) = \int_G XdP$. Then $\nu$ is a vector-measure, continuous with respect to $P$. Let now $\{G_1, \ldots, G_k\}$ be a partition of $\Omega$. Then

$$\sum_{j=1}^k \|\nu(G_j)\| \leq \sum_{j=1}^k \int_{G_j} \|X\|dP = E[\|X\|],$$

so $\|\nu\| \leq \int \|X\|dP$. Then, by the Radon-Nikodŷm property, the desired function exists. For any $\varphi \in V^*$

$$\langle\varphi, \int_G E[X|\mathcal{G}]dP\rangle = \int_G \langle\varphi, X\rangle dP,$$

so that $\langle \varphi, E[X|\mathcal{G}]\rangle = E[\langle \varphi, X\rangle|\mathcal{G}]$ P-a.s.

Since $V$ is separable let $\{\varphi_n\}$ be a sequence in the unit ball of $V^*$ such that $\|v\| = \sup_n |\varphi_n(v)|$ for every $v \in V$. Let now $\Omega_n \in \mathcal{G}$, $P(\Omega_n) = 1$ be such that

$$|\langle \varphi_n, E[X|\mathcal{G}]\rangle| = |E[\langle \varphi_n, X\rangle|\mathcal{G}]| \leq E[\|X\||\mathcal{G}] \quad \text{on } \Omega_n \tag{7}$$

and define $\tilde{\Omega} = \bigcap_n \Omega_n$. Then $P(\tilde{\Omega}) = 1$ and taking supremum on the left hand side of (7) it holds pointwise on $\tilde{\Omega}$ that

$$\|E[X|\mathcal{G}]\| = \sup_n |\langle \varphi_n, E[X|\mathcal{G}]\rangle| \leq E[\|X\||\mathcal{G}],$$

which proves the result.

Finally, to show uniqueness, assume that $\int_A E[X|\mathcal{G}]dP = \int_A ZdP$ for all $A \in \mathcal{G}$. Let $\varphi_n$ be as above, and now let $\Omega_n^0$ have full probability and be such that $\langle \varphi_n, E[X|\mathcal{G}]\rangle = \langle \varphi_n, Z\rangle$ pointwise on $\Omega_n^0$. Since $\{\varphi_n\}$ separates points in $V$, $E[X|\mathcal{G}] = Z$ on $\tilde{\Omega}^0 = \bigcap_n \Omega_n^0$.

□

As noted in the above proof, for any $\varphi \in V^*$ it holds that $\langle \varphi, E[X|\mathcal{G}]\rangle = E[\langle \varphi, X\rangle|\mathcal{G}]$ on some $\Omega_\varphi \in \mathcal{G}$ with $P(\Omega_\varphi) = 1$. It might seem tempting to define the conditional expectation by the above equality, and make a construction similar to the Pettis integral, but as $\Omega_\varphi$ depends on $\varphi \in V^*$, such a construction is difficult.

As the construction of the conditional expectation is a perfect generalisation of the real-valued construction, most properties from the finite-dimensional case, such as the tower property, still hold.

**Lemma 1.14.** *Assume that $X \in L^1(\Omega, \mathcal{F}, P; V)$ has the representation*

$$X = \sum_{n=1}^{\infty} X_n v_n$$

*for two sequences $\{X_n\} \subset L^1(\Omega, \mathcal{F}, P)$ and $\{v_n\} \subset V$ such that $\sum_k E[|X_k|]\|v_k\| < \infty$. Then*

$$E[X|\mathcal{G}] = \sum_{n=1}^{\infty} E[X_n|\mathcal{G}]v_n, \quad P-a.s. \tag{8}$$

*Proof.* This follows directly from noting that

$$\int_G XdP = \sum_{n=1}^{\infty} \left(\int_G X_n dP\right) v_n,$$

since for any $\varphi \in V^*$ it holds that

$$\langle \varphi, \int_G XdP\rangle = \int_G \sum_{n=1}^{\infty} \langle \varphi, v_n\rangle X_n dP = \sum_{n=1}^{\infty} \langle \varphi, v_n\rangle \int_G X_n dP$$

by the dominated convergence theorem. □

Although the lemma is rather trivial, it is included for convenience when discussing the martingale property of Itô integrals in infinite dimensions.

**Vector-valued martingales**

Let $\mathcal{F}_t$, $t \geq 0$ be a filtration on $(\Omega, \mathcal{F}, P)$. The definition of a vector-valued martingale is done precisely as in the finite-dimensional case, i.e. a $V$-valued stochastic process $M$ is called a martingale if

- $M$ is adapted to the filtration $\mathcal{F}_t$,
- $E[\|M(t)\|] < \infty$ for all $t \geq 0$, and
- $E[M(t)|\mathcal{F}_s] = M(s)$ P-a.s.

For a $V$-valued martingale, it follows directly from (6) that the process $t \mapsto \|M(t)\|$ is a submartingale. Indeed

$$\|M(s)\| = \|E[M(t)|\mathcal{F}_s]\| \leq E[\|M(t)\| \,|\mathcal{F}_s]$$

as desired. Also, for a convex function, $f : \mathbb{R}_+ \to \mathbb{R}_+$ the process $t \mapsto f(\|M(t)\|)$ is a real-valued submartingale, since $\|M\|$ is a submartingale. This will be in particular interest when $V = \mathcal{H}$ is a Hilbert space and $f(x) = x^2$.

## 1.3 Hilbert-Schmidt Operators

For an infinite-dimensional separable Hilbert space, it might not hold that $B(\mathcal{H})$, the space of bounded operators, is separable. This leads to trouble when discussing measurability for operator-valued functions. When defining the Itô integral of operator-valued stochastic processes, one also loses the Itô-isometry when using the standard operator norm on $B(\mathcal{H})$. This motivates the following definition.

**Definition 1.15.** *Let $U$ and $\mathcal{H}$ be separable Hilbert-spaces, and $\{f_n\}$ an orthonormal basis for $U$. A linear operator $A : U \to \mathcal{H}$ is called a Hilbert-Schmidt operator if*

$$\sum_{k=1}^{\infty} \|Af_k\|^2 < \infty.$$

If $\{e_k\}$ is an orthonormal basis for $\mathcal{H}$, by Parseval's identity

$$\sum_{k=1}^{\infty} \|Af_k\|^2 = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} |\langle f_k, A^*e_n\rangle|^2 = \sum_{n=1}^{\infty} \|A^*e_n\|^2.$$

So that $A$ is Hilbert-Schmidt if and only if $A^*$ is Hilbert-Schmidt. This also shows that the definition is independent of the choice of orthonormal basis.

Let $L_2(U, \mathcal{H})$ denote the space of all Hilbert-Schmidt operators from $U$ to $\mathcal{H}$, and let

$$\|A\|_2 = \sqrt{\sum_{k=1}^{\infty} \|Af_k\|^2}$$

for $A \in L_2(U, \mathcal{H})$.

**Proposition 1.16.** *The space $L_2(U, \mathcal{H})$ is a separable Hilbert-space with the norm $\|\cdot\|_2$ induced by the inner product*

$$\langle A, B\rangle_2 := \sum_{k=1}^{\infty} \langle Af_k, Bf_k\rangle,$$

*and $L_2(U, \mathcal{H})$ is a subset of the set of compact operators from $U$ to $\mathcal{H}$.*

*Proof.* Let $A \in L_2(U, \mathcal{H})$. When $\{e_n\}$ is an orthonormal basis for $\mathcal{H}$, it holds that for any $u \in U$, $Au = \sum_{n=1}^{\infty} \langle Au, e_n\rangle e_n$. Define $A_m : U \to \mathcal{H}$ by

$$A_m u := \sum_{n=1}^{m} \langle Au, e_n\rangle e_n.$$

Then $A_m$ is a finite rank-operator. It then holds that for a $u \in U$ with $\|u\| \leq 1$, that

$$\|Au - A_m u\|^2 = \sum_{n=m+1}^{\infty} |\langle Au, e_n\rangle|^2 \leq \sum_{n=m+1}^{\infty} \|A^* e_n\|^2 \to 0$$

as $m \to \infty$, since $A^*$ is Hilbert-Schmidt. As the last inequality is independent of $u$, it follows that

$$\|A - A_m\| \to 0$$

as $m \to \infty$. This shows that $A$ is in the closure of the finite-rank operators, hence is compact.
By a similar argument, it follows that

$$\|A\| \leq \|A\|_2.$$

To see that $L_2(U, \mathcal{H})$ is a Hilbert space, let $\{A_j\}$ be a Cauchy sequence in $L_2(U, \mathcal{H})$ with $\|\cdot\|_2$. Since the operator norm is dominated by $\|\cdot\|_2$, $\{A_j\}$ is a Cauchy sequence in $B(U, \mathcal{H})$ with operator norm. Hence, there exists a $A \in B(U, \mathcal{H})$ such that $\|A_j - A\| \to 0$ as $j \to \infty$. Let now $\epsilon > 0$ be given, and $m \in \mathbb{N}$. Since $\{A_j\}$ is Cauchy in the Hilbert-Schmidt norm, for sufficiently large $i$ and $j$,

$$\sum_{k=1}^{m} \|A_i f_k - A_j f_k\|^2 \leq \|A_i - A_j\|_2^2 < \epsilon.$$

Letting $i$ tend to infinity, it follows that

$$\sum_{k=1}^{m} \|Af_k - A_j f_k\|^2 \leq \epsilon$$

Since $\epsilon$ is independent of $m$ and $m$ was arbitrary, it follows that

$$\|A - A_j\|_2^2 \leq \epsilon$$

for sufficiently large $j$, so that $\{A_j\}$ converges in the Hilbert-Schmidt norm. This shows that $L_2(U, \mathcal{H})$ is a Hilbert space.
To see that $L_2(U, \mathcal{H})$ is separable in the Hilbert-Schmidt norm, define the rank-one operator $e_j \otimes f_i$ by

$$(e_j \otimes f_i)u = \langle f_i, u \rangle e_j,$$

which is an orthonormal set in $L_2(U, \mathcal{H})$. If now $A$ is in the orthogonal complement of the set $\{e_j \otimes f_i\}$, it follows that

$$0 = \langle A, e_j \otimes f_i \rangle_2 = \sum_{k=1}^{\infty} \langle Af_k, \langle f_i, f_k \rangle e_j \rangle = \langle Af_i, e_j \rangle$$

for all $i$ and $j$. Since $\{e_j\}$ is an orthonormal basis for $\mathcal{H}$, it follows that $Af_i = 0$. Since this again holds for all $i$ and $\{f_i\}$ is an orthonormal basis for $U$, $A$ must be the zero operator. This shows that $\{e_j \otimes f_i\}$ is an orthonormal basis for $L_2(U, \mathcal{H})$, and it then follows that $L_2(U, \mathcal{H})$ is separable.

□

## 1.4 Itô Integral with respect to Cylindrical Brownian Motion

Based on Example 1.5, this section will make sense of the stochastic integral of Hilbert-space valued functions with respect to Brownian noise.
First, let

$$f : [0, T] \times \Omega \to \mathcal{H}$$

and $B$ be a one-dimensional Brownian motion with usual filtration $\mathcal{F}_t$. As in Example 1.5 it is desirable that

$$\int_0^T f(s) dB(s) = \sum_{n=1}^{\infty} \int_0^T \langle f(s), e_n \rangle dB(s) e_n$$

so that the stochastic integral is an infinite copy of one-dimensional stochastic integrals. This motivates the following definition;

**Definition 1.17.** *A function $f : [0, T] \times \Omega \to \mathcal{H}$ is called Itô-integrable if;*

- $\langle f(t, \cdot), e_n \rangle : \Omega \to \mathbb{R}$ *is $\mathcal{F}_t$-adapted for all $n \in \mathbb{N}$, and*

- $E[\int_0^T |\langle f(s), e_n\rangle|^2 ds] < \infty$, *for all* $n \in \mathbb{N}$.

*Let* $M^2([0,T];\mathcal{H})$ *denote the space of all Itô-integrable functions. For a function* $f \in M^2([0,T];\mathcal{H})$ *define the stochastic integral with respect to* $B$ *as*

$$\int_0^T f(s)dB(s) := \sum_{n=1}^{\infty} \int_0^T \langle f(s), e_n\rangle dB(s) e_n.$$

It is also possible to construct the Itô integral assuming

$$P\left(\int_0^T \|f(s)\|^2 ds < \infty\right) = 1,$$

instead of being square-integrable. This can be done by a standard procedure using localization based on stopping times.

Some of the well-known results about the classical Itô integral remains true for vector valued functions.

**Proposition 1.18.** *The Itô integral has zero expectation, and the Itô isometry holds in the following manner :*

$$E\left[\int_0^T f(s)dB(s)\right] = 0 \quad \textit{(the zero-vector), and}$$

$$E\left[\|\int_0^T f(s)dB(s)\|^2\right] = E\left[\int_0^T \|f(s)\|^2 ds\right]. \tag{9}$$

*Proof.* To see the first equality, let $n \in \mathbb{N}$ be arbitrary. Then

$$\langle E\left[\int_0^T f(s)dB(s)\right], e_n\rangle = E\left[\langle \int_0^T f(s)dB(s), e_n\rangle\right] = E\left[\int_0^T \langle f(s), e_n\rangle dB(s)\right] = 0.$$

Since the vector $E\left[\int_0^T f(s)dB(s)\right]$ is orthogonal to every $e_n$, it must be the zero-vector.

To see (9):

$$E\left[\|\int_0^T f(s)dB(s)\|^2\right] = E\left[\sum_{n=1}^{\infty}\left|\langle \int_0^T f(s)dB(s), e_n\rangle\right|^2\right]$$

$$= \sum_{n=1}^{\infty} E\left[\left|\int_0^T \langle f(s), e_n\rangle dB(s)\right|^2\right] = \sum_{n=1}^{\infty} E\left[\int_0^T |\langle f(s), e_n\rangle|^2 ds\right]$$

$$= E\left[\int_0^T \sum_{n=1}^{\infty} |\langle f(s), e_n\rangle|^2 ds\right] = E\left[\int_0^T \|f(s)\|^2 ds\right].$$

□

As the agenda of this chapter is to translate one-dimensional phenomena to infinite dimensions, this is also done for Brownian noise.

**Definition 1.19 (Cylindrical Brownian motion).** *Let $U$ be a separable Hilbert space with orthonormal basis $\{f_k\}$, and $\{B^k\}$ a sequence of independent one-dimensional Brownian motions. Define*

$$W(t) := \sum_{k=1}^{\infty} B^k(t) f_k, \tag{10}$$

*which is called cylindrical Brownian motion on $U$.*

Notice that the sum in (10) is not convergent. Indeed, for $t > 0$

$$E[\|W(t)\|^2] = E[\sum_{k=1}^{\infty} |B^k(t)|^2] = \sum_{k=1}^{\infty} t = \infty.$$

Nevertheless, the functions that will be integrated with respect to cylindrical Brownian motion will be operator-valued functions. Here the appreciation of the Hilbert-Schmidt operators comes fully into play.
From now on the filtration will be generated by $W$ and $P$-completed, i.e.

$$\mathcal{F}_t := \sigma\{B^k(s) : 0 \le s \le t,\, k \in \mathbb{N}\} \vee \mathcal{N}$$

where $\mathcal{N}$ is the collection of $P$-null sets.

**Definition 1.20.** *Let $\phi \in M^2([0,T]; L_2(U, \mathcal{H}))$. Define the stochastic integral with respect to $W(t)$*

$$\int_0^T \phi(s) dW(s) := \sum_{k=1}^{\infty} \int_0^T \phi(s) f_k dB^k(s).$$

The results of Proposition 1.18 are directly transferred;

**Proposition 1.21.** *The integral has zero expectation*

$$E\left[\int_0^T \phi(s) dW(s)\right] = 0$$

*and by the choice of Hilbert-Schmidt operators, the Itô-isometry still holds*

$$E\left[\left\|\int_0^T \phi(s) dW(s)\right\|^2\right] = E\left[\int_0^T \|\phi(s)\|_2^2 ds\right]. \tag{11}$$

*Proof.* The first equality is obvious by the remark on integration against one-dimensional Brownian motion. To see (11), since the $B^k$s are independent

$$E\left[\| \int_0^T \phi(s)dW(s)\|^2\right] = E\left[\sum_{n=1}^{\infty} |\sum_{k=1}^{\infty} \int_0^T \langle \phi(s)f_k, e_n\rangle dB^k(s)|^2\right]$$

$$= \sum_{n=1}^{\infty} \sum_{k,j=1}^{\infty} E\left[\left(\int_0^T \langle \phi(s)f_k, e_n\rangle dB^k(s)\right)\left(\int_0^T \langle \phi(s)f_j, e_n\rangle dB^j(s)\right)\right]$$

$$= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} E\left[\left(\int_0^T \langle \phi(s)f_k, e_n\rangle dB^k\right)^2\right] = \sum_{k=1}^{\infty} E\left[\int_0^T \|\phi(s)f_k\|^2 ds\right]$$

$$= E\left[\int_0^T \|\phi(s)\|_2^2 ds\right].$$

□

**Lemma 1.22.** *The process $t \mapsto \int_0^t \phi(s)dW(s)$ is a martingale with respect to the filtration, $\{\mathcal{F}_t\}$. Also,*

$$E[\sup_{t\in[0,T]} \| \int_0^t \phi(s)dW(s)\|^2] \leq 4E[\int_0^T \|\phi(s)\|_2^2 ds].$$

*Proof.* In view of (8), this is an easy consequence of the fact that the real-valued Itô integrals are martingales.
Now by Doob's Maximal Inequality (see e.g. [KS98]) applied to the sub-martingale $M(t) := \| \int_0^t \phi(s)dW(s)\|$ it follows that

$$E[\sup_{t\in[0,T]} M(t)^2] \leq 4E[M(T)^2] = 4E[\int_0^T \|\phi(s)\|_2^2 ds]$$

by the Itô-isometry. □

## 1.5 Differentiation

The definition of the derivative for a vector valued function will be exactly the same as for the one-dimensional case.

**Definition 1.23.** *Let $V$ be a Banach space, $\Lambda \subset \mathbb{R}$ be an open interval, and $f : \Lambda \to V$. The function will be called differentiable at a point $t \in \Lambda$ if there exists vector $y \in V$ such that*

$$\|\frac{1}{h}(f(t+h) - f(t)) - y\| \to 0$$

*as $h \to 0$. Denote the derivative of $f$ at $t$ by $f'(t)$. If the function is differentiable at all points in $\Lambda$, it is called differentiable, and the function $f' : t \mapsto f'(t)$ is called the derivative of $f$. Iterating this procedure $n$ times gives the $n$-th derivative, denoted $f^{(n)}$. The space of $n$-times differentiable functions from $\Lambda$ to $V$ will be denoted $C^n(\Lambda; V)$.*

It is clear that a differentiable function has to be continuous, but as in the usual sense a continuous function is not necessarily differentiable.

**Proposition 1.24.** *If $f \in C^1(\Lambda; V)$ and $\varphi \in V^*$, the function $\varphi \circ f : \Lambda \to \mathbb{R}$ is differentiable in the usual sense, and*

$$(\varphi \circ f)'(t) = \varphi \circ f'(t).$$

*Proof.* By the linearity and continuity of $\varphi$,

$$\lim_{h \to 0} \frac{\varphi(f(t+h)) - \varphi(f(t))}{h} = \varphi \left( \lim_{h \to 0} \frac{f(t+h) - f(t)}{h} \right)$$

which gives the desired result. □

**Proposition 1.25 (Fundamental theorem of calculus).** *Let $f \in C^1(\Lambda, V)$ and $s, t \in \Lambda$, with $s < t$. Then*

$$f(t) = f(s) + \int_s^t f'(u) du.$$

*Proof.* Let $\varphi \in V^*$, and let $g := \varphi \circ f$. From Proposition 1.24 $g \in C^1(\Lambda)$ and by the Fundamental theorem of calculus $g(t) - g(s) = \int_s^t g'(u) du$ and $g' = \varphi \circ f'$, so

$$\begin{aligned} {}_{V^*}\langle f(t), \varphi \rangle_V - {}_{V^*}\langle f(s), \varphi \rangle_V &= {}_{V^*}\langle f(t) - f(s), \varphi \rangle_V \\ &= \int_s^t {}_{V^*}\langle f'(u), \varphi \rangle_V du = {}_{V^*}\langle \int_s^t f'(u) du, \varphi \rangle_V. \end{aligned}$$

Since $\varphi \in V^*$ was arbitrary, the result follows.

□

**Proposition 1.26.** *Assume that $\mathcal{H}$ is a Hilbert space and $f, g \in C^1(\Lambda, \mathcal{H})$. Then the function $\langle f(\cdot), g(\cdot) \rangle : \Lambda \to \mathbb{R}$ is in $C^1(\Lambda)$ and*

$$(\langle f(t), g(t) \rangle)' = \langle f'(t), g(t) \rangle + \langle f(t), g'(t) \rangle.$$

*In particular, $\|f(\cdot)\|^2 \in C^1(\Lambda)$ and*

$$\left( \|f(t)\|^2 \right)' = 2 \langle f'(t), f(t) \rangle. \tag{12}$$

*Proof.* Writing

$$\begin{aligned} &\frac{1}{h}(\langle f(t+h), g(t+h) \rangle - \langle f(t), g(t) \rangle) \\ &= \frac{1}{h}(\langle f(t+h), g(t+h) \rangle - \langle f(t), g(t+h) \rangle + \langle f(t), g(t+h) \rangle - \langle f(t), g(t) \rangle) \\ &= \langle \frac{1}{h}(f(t+h) - f(t)), g(t+h) \rangle + \langle f(t), \frac{1}{h}(g(t+h) - g(t)) \rangle \end{aligned}$$

and using Proposition 1.24, the result follows. □

## 1.6 Strongly Continuous Semi-groups

**Definition 1.27.** *Let $V$ be a Banach space. A family $\{S(t)\}_{t\geq 0}$ of operators in $B(V)$ is called a semi-group (of operators) if*

- $S(t)S(s) = S(t+s)$ ,
- $S(0) = I$ .

*A semi-group for which the map $t \mapsto S(t)$ is continuous when $B(V)$ is equipped with the strong operator topology, is called a strongly continuous semi-group. This means that the map $t \mapsto S(t)x$ is continuously $V$-valued for every $x \in V$.*

Later on, it will be desirable to be able to bound $\|S(t)\|$ independently of $t$. When dealing with a finite time-horizon, this is always possible.

**Lemma 1.28.** *For a strongly continuous semi-group $\{S(t)\}_{t\in[0,T]}$ where $T > 0$ is fixed,*

$$\sup_{t\in[0,T]} \|S(t)\| < \infty.$$

*Proof.* Since $[0,T]$ is compact and $t \mapsto S(t)x$ is continuous, the set

$$\{S(t)x | t \in [0,T]\}$$

is compact, hence bounded in $V$. By the Banach-Steinhaus theorem, it follows that the set

$$\{\|S(t)\| \,|\, t \in [0,T]\}$$

is bounded. □

**Example 1.29 (Left-translation semi-group).** *Let $V = C_b(\mathbb{R})$ with supremum-norm, and define $S(t) \in B(V)$ by $(S(t)f)(x) = f(x+t)$. Then $\{S(t)\}_{t\geq 0}$ is a semi-group and is also strongly continuous.*

**Example 1.30.** *Let $B(t)$ be a Brownian motion on $\mathbb{R}^n$, and let*

$$b : \mathbb{R}^n \to \mathbb{R}^n$$

$$\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times n}$$

*be such that there exists a solution to the stochastic differential equation*

$$\begin{cases} dX(t) = b(X(t))dt + \sigma(X(t))dB(t) \\ X(0) = x \end{cases}$$

*for any $x \in \mathbb{R}^n$. Denote its solution (which depends on $x$) by $X^x(t)$.*

*Let $V = B^\infty(\mathbb{R}^n)$, and define $S(t) : B^\infty(\mathbb{R}^n) \to B^\infty(\mathbb{R}^n)$ by*

$$(S(t)f)(x) = E\left[f(X^x(t))\right].$$

*By the linearity of the expectation, $S(t)$ is a linear operator, and since*

$$|E\left[f(X^x(t))\right]| \leq E\left[|f(X^x(t))|\right] \leq E[\|f\|_\infty] = \|f\|_\infty$$

*$S(t)$ is indeed in $B(V)$, and $\|S(t)\| \leq 1$. By the Markov-property of the diffusion $X^x(t)$, it follows that*

$$(S(t)S(s)f)(x) = S(t)\left(E^{\cdot}\left[f(X^x(s))\right]\right)(x) = E\left[E\left[f(X^{X^x(t)}(s))\right]\right]$$

$$= E\left[E\left[f(X^x(t+s))|\mathcal{F}_t\right]\right] = E\left[f(X^x(t+s))\right] = (S(t+s)f)(x)$$

*so that $S(t)S(s) = S(t+s)$.*

*When restricted to $C_0^2(\mathbb{R}^n)$, the semi-group is strongly continuous. Indeed, by Dynkin's formula (see [Øks05]), for $f \in C_0^2(\mathbb{R}^n)$*

$$E\left[f(X^x(t))\right] = f(x) + E\left[\int_0^t Af(X^x(s))ds\right],$$

*where*

$$A = \sum_{i=1}^n b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n (\sigma\sigma^T)_{i,j}(x)\frac{\partial^2}{\partial x_i \partial x_j},$$

*and hence*

$$|S(t)f(x) - f(x)| \leq \int_0^t E\left[|Af(X^x(s))|\right] ds \to 0$$

*as $t \to 0$ for all $x \in \mathbb{R}^n$, and so*

$$\|S(t)f - f\|_\infty \to 0.$$

Notice that the supremum-norm is not the canonical norm on $C_0^2(\mathbb{R}^n)$, so that the above examples does not show that $S(t)$ is strongly continuous on $B^\infty(\mathbb{R}^n)$. Rigorous information on this subject can be found in [MFT94].

**Definition 1.31.** *Let $S(t)$ be a strongly continuous semi-group of operators on a Banach space $V$, and let*

$$\mathcal{D}(A) := \left\{v \in V : \lim_{h \to 0} \frac{S(h)v - v}{h} \text{ exists in } V\right\}.$$

*Define $A : \mathcal{D}(A) \to V$ by*

$$Av = \lim_{h \to 0} \frac{S(h)v - v}{h}.$$

Since

$$\frac{S(h)(\alpha v + \beta u) - (\alpha v + \beta u)}{h} = \alpha \frac{S(h)v - v}{h} + \beta \frac{S(h)u - u}{h}$$

it follows that $\mathcal{D}(A)$ is a linear subspace of $V$ and that $A(\alpha v + \beta u) = \alpha Av + \beta Au$ so $A$ is a linear operator. The following examples will show that the operator $A$ is not continuous in general.

**Example 1.32.** *For the right-translation semi-group in Example 1.29, it is immediate that*

$$C^1(\mathbb{R}) \subset \mathcal{D}(A)$$

*and that $A = \frac{d}{dx}$, on $C^1(\mathbb{R})$.*

**Example 1.33.** *In Example 1.30, again by Dynkin's formula, $C_0^2(\mathbb{R}^n) \subset \mathcal{D}(A)$, and for a function $f \in C_0^2(\mathbb{R}^n)$, by the Fundamental Theorem of Calculus*

$$\frac{1}{h}(S(h)f(x) - f(x)) = \frac{1}{h}\int_0^h E[Af(X^x(s))]ds \to E[Af(X^x(0))] = Af(x),$$

*where $A$ is as before.*

**Proposition 1.34.** *If $x \in \mathcal{D}(A)$, then for all $t \geq 0$, $S(t)x \in \mathcal{D}(A)$. In this case the function $t \mapsto S(t)x$ is differentiable (differentiable from the right at $t = 0$), and*

$$\frac{d}{dt}S(t)x = S(t)Ax = AS(t)x.$$

*Proof.* Let $t > 0$. By the continuity of $S(t)$ and definition of $\frac{d}{dt}S(t)x$,

$$\lim_{h\to 0} \frac{S(t+h)x - S(t)x}{h} = \lim_{h\to 0} \frac{S(t)(S(h)x - x)}{h}$$

$$= S(t) \lim_{h\to 0} \frac{S(h)x - x}{h} = S(t)Ax.$$

It is also clear that $AS(t) = S(t)A$ on $\mathcal{D}(A)$. □

This result will be of particular interest when considering $V$-valued differential equations of the form

$$\begin{aligned} \tfrac{du}{dt} &= Au \\ u(0) &= x \end{aligned} \tag{13}$$

where $A$ is the generator of a strongly continuous semi-group, and $x \in \mathcal{D}(A)$. Proposition 1.34 states that the function $u(t) = S(t)x$ is a solution to (13). More can be said, and in [Bob05] uniqueness is proved.

**Lemma 1.35.** *There exists a unique solution to (13) given by $u(t) = S(t)x$.*

# 2 Some Theory from Partial Differential Equations

As noted in Chapter 1.6, it is possible to consider a partial differential equation as an ordinary differential equation consisting of vector-valued functions. Unfortunately, differentiation is not a continuous operator on e.g. $L^2(\mathbb{R})$. One way of overcoming this problem is addressed via strongly continuous semi-groups. Another way is to consider variational solutions, as will be presented here.

## 2.1 Gelfand Triples

Let $\mathcal{H}$ be a separable Hilbert-space and $V$ a reflexive Banach-space such that the embedding $V \hookrightarrow \mathcal{H}$ is continuous and dense, i.e. there exists a $J \in B(V, \mathcal{H})$ such that $\ker J = \{0\}$ and $J(V)$ is dense in $\mathcal{H}$.

**Proposition 2.1.** *Let $V$ and $\mathcal{H}$ be as above. Then $\mathcal{H}^* \hookrightarrow V^*$ is continuous and dense.*

*Proof.* Define the map $J^* : \mathcal{H}^* \to V^*$ by ${}_{V^*}\langle J^*(\varphi), v\rangle_V = \langle \varphi, J(v)\rangle$ for all $\varphi \in \mathcal{H}^*$ and $v \in V$. Then $\ker J^* = \{0\}$. Indeed, assume that $\langle \varphi, J(v)\rangle = 0$ for all $v \in V$. Since $J(V)$ is dense in $\mathcal{H}$, $\varphi = 0$. By the closed graph theorem, it follows that $J^* \in B(\mathcal{H}^*, V^*)$.
Assume that $J^*(\mathcal{H}^*)$ is not dense in $V^*$ and consider the closure $J^*(\mathcal{H}^*)^-$. By the Hahn-Banach theorem, we may choose a functional $\psi \in V^{**}$ such that $\|\psi\| = 1$ and $\psi|_{J^*(\mathcal{H}^*)^-} = 0$. Now, since $V^{**} = V$ and all Hilbert spaces are reflexive, it follows that the iterated dual $J^{**}$ is equal to $J$. Indeed,

$$\langle \varphi, J^{**}(v)\rangle = {}_{V^*}\langle J^*(\varphi), v\rangle_V = \langle \varphi, J(v)\rangle.$$

Now, the choice of $\psi$ is such that $\varphi \in \ker J^{**} = \ker J = \{0\}$ which is a contradiction.

∎

The embedding $V \hookrightarrow \mathcal{H}$ will be written $V \subset \mathcal{H}$ and the map $J$ will be dropped in the notation. The examples that follow will justify this notation. Identifying $\mathcal{H}$ with its dual via the Riesz identification it follows that

$$V \subset \mathcal{H} \subset V^*$$

continuously and densely. The triple $(V, \mathcal{H}, V^*)$ is called a *Gelfand triple.* By the definition of the embeddings it also holds that for a $h \in \mathcal{H}$, when considered as an element of $V$,

$${}_{V^*}\langle h, v\rangle_V = \langle h, v\rangle$$

for all $v \in V$ when considered as an element of $\mathcal{H}$. In the remainder, ${}_{V^*}\langle \cdot, \cdot\rangle_V$ will denote the dual pairing between $V$ and $V^*$ with norms $\|\cdot\|_V$ and $\|\cdot\|_{V^*}$, respectively. The inner product on $\mathcal{H}$ will simply be denoted by $\langle \cdot, \cdot\rangle$ and the induced norm by $\|\cdot\|$.

**Example 2.2.** *Let $p > 2$, and $\Lambda \subset \mathbb{R}^n$ be open, with $\lambda(\Lambda) < \infty$ where $\lambda$ is the Lebesgue measure on $\mathbb{R}^n$. Then $L^p(\Lambda) \subset L^2(\Lambda) \subset L^{p/(p-1)}(\Lambda)$ is a Gelfand triple.*

*Proof.* For a function $u \in L^p(\Lambda)$, we have by the Hölder inequality

$$\int_\Lambda |u|^2 d\lambda \leq (\lambda(\Lambda))^{(p-2)/p} \left(\int_\Lambda |u|^p d\lambda\right)^{2/p} < \infty,$$

so that $u \in L^2(\Lambda)$, and the embedding is just the identity map from $L^p(\Lambda)$ to $L^2(\Lambda)$. This justifies the notation $L^p(\Lambda) \subset L^2(\Lambda)$. Since $\lambda(\Lambda) < \infty$, all step-functions on $\Lambda$ are in both $L^p(\Lambda)$ and $L^2(\Lambda)$. It then follows that $L^p(\Lambda)$ is dense in $L^2(\Lambda)$. Finally, since $(L^p(\Lambda))^* = L^{p/(p-1)}(\Lambda)$ the result follows. $\square$

To get some more interesting examples of Gelfand triples and useful modeling spaces for solutions of SPDE's, it is convenient to introduce the notion of Sobolev spaces.

## 2.2 Weak Derivatives

Let $\Lambda$ be a open subset of $\mathbb{R}^n$, let $u \in C^1(\Lambda)$ and $\phi \in C_c^\infty(\Lambda)$. By integration by parts, it follows that

$$\int_\Lambda u \frac{\partial \phi}{\partial x_i} d\lambda = - \int_\Lambda \phi \frac{\partial u}{\partial x_i} d\lambda$$

More generally, let $\mathbb{N}^n$ be equipped with the one-norm, $|\cdot|_1$, and define $D^\alpha := \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. For $u \in C^k(\Lambda)$ and $\phi \in C_c^\infty(\Lambda)$, iterating the integration by parts gives

$$\int_\Lambda u D^\alpha \phi d\lambda = (-1)^{|\alpha|_1} \int_\Lambda \phi D^\alpha u d\lambda$$

for $|\alpha|_1 \leq k$. This motivates the following definition :

**Definition 2.3.** *A function $u \in L^1_{loc}(\Lambda)$, $\alpha \in \mathbb{N}^n$ has a weak $\alpha$-th derivative, denoted $D^\alpha u$, provided*

$$\int_\Lambda u D^\alpha \phi d\lambda = (-1)^{|\alpha|_1} \int_\Lambda \phi D^\alpha u d\lambda$$

*for all $\phi \in C_c^\infty(\Lambda)$.*

Since the equality is to be for all $\phi \in C_c^\infty(\Lambda)$, the weak derivative, if it exists, it is uniquely defined up to a set of Lebesgue measure zero. By the above discussion, this clearly extends the notion of differentiability.

## 2.3 Sobolev Spaces

**Definition 2.4.** *Let $1 \leq p < \infty$. Define $W^{k,p}(\Lambda)$ to be the space of all $u \in L^1_{loc}(\Lambda)$ such that its $\alpha$-th weak derivative $D^\alpha u$ exists, and $D^\alpha u \in L^p$ for all $|\alpha|_1 \leq k$. Define the norm $\|\cdot\|_{k,p}$ on $W^{k,p}(\Lambda)$ by*

$$\|u\|_{k,p} = \left( \int_\Lambda (|u|^p + \sum_{|\alpha|_1 \leq k} |D^\alpha u|^p) d\lambda \right)^{1/p}.$$

*The space $W^{k,p}(\Lambda)$ with $\|\cdot\|_{k,p}$ is then a Banach-space, and is called the Sobolev space of order $k$ in $L^p(\Lambda)$.*

When $p = 2$ one writes $H^k(\Lambda) := W^{k,2}(\Lambda)$ and $\|\cdot\|_{H^k} := \|\cdot\|_{k,2}$. Clearly, when equipped with the inner product

$$\langle f, g\rangle_{H^k} = \int_\Lambda fg + \sum_{|\alpha|_1 \leq k} (D^\alpha f)(D^\alpha g)\, d\lambda$$

this becomes a Hilbert space.

**Definition 2.5.** *Denote by $W^{k,p}_0(\Lambda)$ the closure of $C^\infty_c(\Lambda)$ in $W^{k,p}(\Lambda)$, i.e.*

$$W^{k,p}_0(\Lambda) = (C^\infty_c(\Lambda))^{-\|\cdot\|_{k,p}}.$$

Similarly, define $H^k_0(\Lambda) := W^{k,2}_0(\Lambda)$. $W^{k,p}_0(\Lambda)$ is to be thought of as the functions in $W^{k,p}(\Lambda)$ which vanish near the boundary of $\Lambda$.

**Example 2.6.** *Let $\Lambda \subset \mathbb{R}^n$, now possibly with infinite measure. Define $H^{-1}(\Lambda) := \left(H^1_0(\Lambda)\right)^*$. Then $(H^1_0(\Lambda), L^2(\Lambda), H^{-1}(\Lambda))$ is a Gelfand triple.*

This example of a Gelfand triple has some useful properties: Let $\Delta := \sum_{i=1}^n \frac{\partial^2}{\partial^2 x_i}$ be the Laplace operator. With $\mathcal{D}(\Delta) = C^2(\Lambda)$ and $\Delta$ regarded as an operator on $L^2(\Lambda)$, it is not continuous. But defining $\Delta$ as an operator from $H^1_0(\Lambda)$ into $H^{-1}(\Lambda)$, it becomes a continuous operator. To see this, let $\varphi, \psi \in C^\infty_c(\Lambda)$. Then, by integration by parts gives

$$|{}_{H^{-1}}\langle \Delta\varphi, \psi\rangle_{H^1_0}| = |\int_\Lambda (\Delta\varphi)\, \psi\, d\lambda| = |-\int_\Lambda (\nabla\varphi)\cdot(\nabla\psi) d\lambda|$$

$$\leq \left(\int_\Lambda |\nabla\varphi|^2 d\lambda\right)^{1/2} \left(\int_\Lambda |\nabla\psi|^2 d\lambda\right)^{1/2} \leq \|\varphi\|_{H^1}\|\psi\|_{H^1},$$

where the second last inequality follows from Hölders inequality. It then follows that $\Delta\varphi$ is continuous on $C^\infty_c(\Lambda)$. Since $C^\infty_c(\Lambda)$ is dense (by definition) in $H_0^1(\Lambda)$, $\Delta\varphi$ can be extended to a continuous linear functional on $H_0^1(\Lambda)$ satisfying

$$\|\Delta\varphi\|_{H^{-1}} \leq \|\varphi\|_{H^1}$$

on $C_c^\infty(\Lambda)$. Using again that $C_c^\infty(\Lambda)$ is dense in $H^1(\Lambda)$, $\Delta$ can be uniquely extended to a linear operator (still denoted by $\Delta$)

$$\Delta : H_0^1(\Lambda) \to H^{-1}(\Lambda)$$

which is continuous, and $\|\Delta\| \leq 1$.

## 2.4 Variational Solutions of Partial Differential Equations

Let $V \subset \mathcal{H} \subset V^*$ be a Gelfand-triple. Consider the equation

$$\begin{cases} \frac{du(t)}{dt} = & Au(t) + f(t) \\ u(0) = & u_0 \in \mathcal{H}, \end{cases} \tag{14}$$

where $A$ is linear operator from $V$ to $V^*$ and $f \in L^2([0,T]; V^*)$.

**Theorem 2.7.** *Assume that $A$ is continuous and that there exist constants $\lambda \geq 0$ and $\alpha > 0$ such that*

$$2\,{}_{V^*}\langle A\varphi, \varphi\rangle_V \ \leq \lambda\|\varphi\|^2 - \alpha\|\varphi\|_V^2 \tag{15}$$

*for every $\varphi \in V$.*

*Then there exists a unique continuously $\mathcal{H}$-valued function $u \in L^2([0,T]; V)$ such that $u$ satisfies (14).*

*Proof.* As $V$ is dense in $\mathcal{H}$, choose an orthonormal basis, $\{e_j : j \in \mathbb{N}\}$ of $\mathcal{H}$ such that $\operatorname{span}\{e_j : j \in \mathbb{N}\}$ is dense in V.
Let $n \in \mathbb{N}$ and for $1 \leq j \leq n$ define $u_{j,n}$ to be the (real-valued) solution of

$$\frac{du_{j,n}(t)}{dt} = \sum_{i=1}^n u_{i,n}(t)_{V^*}\langle Ae_i, e_j\rangle_V + {}_{V^*}\langle f(t), e_j\rangle_V$$

$$u_{j,n}(0) = \langle u_0, e_j\rangle.$$

Define $u_n(t) = \sum_{j=1}^n u_{j,n}(t)e_j$. Then $u_n$ satisfies

$$\langle\frac{du_n(t)}{dt}, e_j\rangle = {}_{V^*}\langle Au_n(t), e_j\rangle_V + {}_{V^*}\langle f(t), e_j\rangle_V$$

$$u_n(0) = \sum_{j=1}^n \langle u_0, e_j\rangle e_j$$

for every $j \in \mathbb{N}$, so that the first line above reads

$$\frac{du_n(t)}{dt} = Au_n(t) + f(t).$$

By construction, $u_n$ is $V$-valued, and can thus be regarded as $\mathcal{H}$-valued. By the chain rule (12)

$$\frac{d\|u_n(t)\|^2}{dt} = 2\left\langle \frac{du_n(t)}{dt}, u_n(t) \right\rangle = 2_{V^*}\langle Au_n(t), u_n(t)\rangle_V + 2_{V^*}\langle f(t), u_n(t)\rangle_V.$$

By condition (15),

$$\|u_n(t)\|^2 = \|u_n(0)\|^2 + \int_0^t 2_{V^*}\langle Au_n(s), u_n(s)\rangle_V + 2_{V^*}\langle f(s), u_n(s)\rangle_V ds$$

$$\leq \|u_0\|^2 + \int_0^t \lambda\|u_n(s)\|^2 - \alpha\|u_n(s)\|_V^2 + 2\|f(s)\|_{V^*}\|u_n(s)\|_V ds.$$

For positive real numbers a, b and $\beta$, it holds that $2ab = 2\left(\frac{a}{\sqrt{\beta}}\right)\left(\sqrt{\beta}b\right) \leq \frac{a^2}{\beta} + \beta b^2$. Putting $a = \|f(s)\|_{V^*}$, $b = \|u_n(s)\|_V$ the above is dominated by

$$\|u_0\|^2 + \int_0^t \lambda\|u_n(s)\|^2 - (\alpha - \beta)\|u_n(s)\|_V^2 + \beta^{-1}\|f(s)\|_{V^*}^2 ds.$$

Choosing $\beta = \alpha/2$ gives

$$\|u_n(t)\|^2 + \frac{1}{2}\int_0^t \|u_n(s)\|_V^2 ds \leq \|u_0\|^2 + \int_0^t \lambda\|u_n(s)\|^2 + 2\alpha^{-1}\|f(s)\|_{V^*}^2 ds. \tag{16}$$

Also, by Gronwall's inequality, we have

$$\sup_{t\in[0,T]} \|u_n(t)\|^2 \leq \left(\|u_0\|^2 + 2\alpha^{-1}\int_0^T \|f(s)\|_{V^*}^2 ds\right) e^{\lambda T}.$$

Using this in (16) it also holds that

$$\int_0^T \|u_n(s)\|_V^2 ds \leq K$$

for some constant K which depends on $\alpha, \beta, f$ and $T$, but not on $n$. This gives that $\{u_n\}$ is a bounded sequence in $L^2([0,T]; V)$, and so there exists a $u$ in $L^2([0,T]; V)$ and a subsequence (still indexed by $n$) such that

$$u_n \to u$$

in the weak topology on $L^2([0,T]; V)$. To see that $u$ is the desired solution, let $\varphi \in L^2([0,T]; V)$. Then by the definition of weak convergence,

$$\int_0^T {}_{V^*}\langle \varphi(t), u(t)\rangle_V dt = \lim_{n\to\infty} \int_0^T {}_{V^*}\langle \varphi(t), u_n(t)\rangle_V dt.$$

Now, for every $n \in \mathbb{N}$

$$\int_0^T \left( {}_{V^*}\langle \varphi(t), u_n(0) \rangle_V + \int_0^t {}_{V^*}\langle Au_n(s), \varphi(t) \rangle_V + {}_{V^*}\langle f(s), \varphi(t) \rangle_V ds \right) dt$$

$$= \int_0^T {}_{V^*}\langle \varphi(t), u_n(0) \rangle_V dt + \int_0^T {}_{V^*}\langle Au_n(s), \int_s^T \varphi(t)dt \rangle_V + {}_{V^*}\langle f(s), \int_s^T \varphi(t)dt \rangle_V ds,$$

which converges to

$$\int_0^T {}_{V^*}\langle \varphi(t), u_0 \rangle_V dt + \int_0^T {}_{V^*}\langle Au(s), \int_s^T \varphi(t)dt \rangle_V + {}_{V^*}\langle f(s), \int_s^T \varphi(t)dt \rangle_V ds$$

$$= \int_0^T \left( {}_{V^*}\langle \varphi(t), u_0 \rangle_V + \int_0^t {}_{V^*}\langle Au(s), \varphi(t) \rangle_V + {}_{V^*}\langle f(s), \varphi(t) \rangle_V ds \right) dt.$$

as $n \to \infty$. Let now $\varphi_0 \in L^\infty[0,T]$ and $j \in \mathbb{N}$, and replace $\varphi$ by $\varphi_0(t)e_j$. This gives that

$$\langle u(t), e_j \rangle = \langle u_0, e_j \rangle + \int_0^t {}_{V^*}\langle Au(s), e_j \rangle_V + {}_{V^*}\langle f(s), e_j \rangle_V ds$$

for every $j$, so that in fact

$$u(t) = u_0 + \int_0^t Au(s) + f(s)ds$$

in $\mathcal{H}$ as desired.

To see that $u$ is continuously $\mathcal{H}$-valued let $r \le t$ and look at the $\mathcal{H}$-valued function $t \mapsto u(t) - u(r) = \int_r^t Au(s) + f(s)ds$. Then

$$\|u(t) - u(r)\|^2 = 2\int_r^t \langle Au(s), u(s) - u(r) \rangle + \langle f(s), u(s) - u(r) \rangle ds$$

which converges to 0 as $r \to t$ since $u \in L^2([0,T];V)$ and $f \in L^2([0,T];V^*)$. Finally, to show uniqueness, assume that both $u_1$ and $u_2$ solve (14). Then $y := u_1 - u_2$ satisfy

$$\frac{dy(t)}{dt} = Ay(t),$$
$$y(0) = 0.$$

Again, by the chain rule

$$\|y(t)\|^2 = \int_0^t 2_{V^*}\langle Ay(s), y(s) \rangle_V ds$$

$$\le \lambda \int_0^t \|y(s)\|^2 ds - \alpha \int_0^t \|y(s)\|_V^2 ds \le \lambda \int_0^t \|y(s)\|^2 ds$$

so by Gronwall's inequality

$$y(t) = 0$$

for all $t \in [0,T]$. □

**Example 2.8 (Heat equation).** *There exists a unique variational solution to*

$$\frac{du}{dt} = \Delta u(t) + f(t)$$

$$u(0) = u_0 \in L^2(\Lambda)$$

*on the Gelfand triple* $H_0^1(\Lambda) \subset L^2(\Lambda) \subset H^{-1}(\Lambda)$, *where* $f \in L^2([0,T]; H^{-1}(\Lambda))$

*Proof.* It was noted in the discussion following Example (2.6) that $\Delta$ is continuous when regarded as a map from $H_0^1(\Lambda)$ to $H^{-1}(\Lambda)$. To see that $\Delta$ satisfies (15), let $u \in H_0^1(\Lambda)$, and just consider

$$\langle \Delta u, u \rangle = -\int_\Lambda |\nabla u|^2 d\lambda = \|u\|^2 - \|u\|_{1,2}^2.$$

□

# 3 Stochastic Equations in Infinite Dimensions

The first chapter made it possible to generalize n-dimensional Itô-processes to infinite-dimensional Itô-processes. The next step is then to consider stochastic differential equations in infinite dimensions, and this will be done in this section.

For the remainder of this chapter it will be assumed

- $V \subset \mathcal{H} \subset V^*$ is a Gelfand triple.
- $W$ is a cylindrical Brownian motion defined on another separable Hilbert-space $U$ with orthonormal basis $\{f_k\}$.
- $(\Omega, \mathcal{F}, P)$ is a complete probability space with $\mathcal{F}_t$ the usual filtration generated by $W$, i.e.

$$\mathcal{F}_t := \sigma\{B^k(s) : 0 \leq s \leq t \,,\, k \in \mathbb{N}\} \vee \mathcal{N}$$

  where $\mathcal{N}$ is the collection of $P$-null sets.
- $T > 0$ denotes the finite time-horizon. Although initially fixed, it will be allowed to vary later on, in order to construct some contraction-mappings.

For notational convenience, for a Banach space $V$, introduce

$$M^p([0,T];V) := \{f \in L^p([0,T] \times \Omega; V) \,|\, f \text{ is adapted to } \mathcal{F}_t\} \,.$$

To see that this is a Banach-space it is sufficient to note that it is a closed subspace of $L^p([0,T] \times \Omega; V)$. This follows immediately by noting that limits of measurable functions is again measurable, see [PR07].

## 3.1 Itô's Formula

In the finite-dimensional case, the Itô-formula is essential for showing existence of solutions to stochastic differential equations. Below we present a variation of the Itô-formula. It is not nearly as strong as for the finite-dimensional case, but still it extends an important way of using the formula. The proof presented here is only a sketch and is based on the proof from [PR07].

**Theorem 3.1.** *Let $\alpha > 1$ and assume*

$$\begin{aligned} &X_0 \in L^2(\Omega, \mathcal{F}_0, P; \mathcal{H}), \\ &Y \in M^{\alpha/(\alpha-1)}([0,T]; V^*), \\ &Z \in M^2([0,T]; L_2(U; \mathcal{H})). \end{aligned}$$

*Define the $V^*$-valued, adapted continuous process*

$$X(t) = X_0 + \int_0^t Y(s)ds + \int_0^t Z(s)dW(s).$$

*If $X \in M^\alpha([0,T];V)$, then $X(\cdot)$ is in fact an $\mathcal{H}$-valued, adapted continuous process with*

$$E\left[\sup_{t\in[0,T]} \|X(t)\|^2\right] < \infty.$$

*In addition, the real-valued process $\|X(\cdot)\|^2$ has the form*

$$\|X(t)\|^2 = \|X_0\|^2 + \int_0^t 2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 ds + \int_0^t 2\langle X(s), Z(s)dW(s)\rangle$$

*Proof.* Since both $Y$ and $Z$ are adapted, it follows that the processes $t \mapsto \int_0^t Y(s)ds$ and $t \mapsto \int_0^t Z(s)dW(s)$ are adapted processes. Since they also are continuous, $X(t)$ is a continuously $V^*$-valued.

The first part of the proof will be to show that $E[\sup_t \|X(t)\|^2] < \infty$.

Let $t > s$ be such that $X(t)$ and $X(s)$ are in $V$. By calculation, it follows that

$$\|X(t)\|^2 - \|X(s)\|^2 = 2\int_s^t {}_{V^*}\langle Y(r), X(t)\rangle_V dr + 2\langle X(s), \int_s^t Z(r)dW(r)\rangle$$

$$+\|\int_s^t Z(r)dW(r)\|^2 - \|X(t) - X(s) - \int_s^t Z(r)dW(r)\|^2$$

Let now $I_l$ be a sequence of partitions, $I_l = \{0 = t_0^l < t_1^l < \ldots < t_{k^l}^l = t\}$ such that

- $X(t_i^l) \in V$ for all $i = 0, 1, \cdots, k_l$ and $l \in \mathbb{N}$,
- $I_l \subset I_{l+1}$ for every $l$, and $\sup_i |t_{i+1}^l - t_i^l| \to 0$ as $l \to \infty$, and
- the processes

  $$\bar{X}^l := \sum_{i=2}^{k_l} X(t_{i-1}^l)\chi_{[t_{i-1}^l, t_i^l)}$$

  and

  $$\tilde{X}^l := \sum_{i=1}^{k_l - 1} X(t_i^l)\chi_{[t_{i-1}^l, t_i^l)}$$

  both converge to $X$ in $L^\alpha([0,t]\times\Omega; V)$ as $l \to \infty$

Notice that $\bar{X}^l$ is adapted to $\mathcal{F}_t$ while $\tilde{X}^l$ is not.

For a fixed $l \in \mathbb{N}$, using the above formula for the partition points in $I_l$ it then follows that

$$\|X(t)\|^2 - \|X_0\|^2 = \sum_{j=0}^{k_l} \left( \|X(t_{j+1}^l)\|^2 - \|X(t_j^l)\|^2 \right) \tag{17}$$

$$= 2\int_0^t {}_{V^*}\langle Y(s), \tilde{X}^l(s)\rangle_V ds + 2\int_0^t \langle \bar{X}^l(s), Z(s)dW(s)\rangle \tag{18}$$

$$+\sum_{j=0}^{k_l} \| \int_{t_j^l}^{t_{j+1}^l} Z(s)dW(s)\|^2 - \|X(t_{j+1}^l) - X(t_j^l) - \int_{t_j^l}^{t_{j+1}^l} Z(s)dW(s)\|^2 \tag{19}$$

$$\leq 2\int_0^t {}_{V^*}\langle Y(s), \tilde{X}^l(s)\rangle_V ds + 2\int_0^t \langle \bar{X}^l(s), Z(s)dW(s)\rangle + \sum_{j=0}^{k_l} \| \int_{t_j^l}^{t_{j+1}^l} Z(s)dW(s)\|^2$$

Using that $\bar{X}^l \to X$ and $\tilde{X}^l \to X$ as $l \to \infty$ one can get

- $E[\sup_{t\in I_l} | \int_0^t {}_{V^*}\langle Y(s), \tilde{X}^l(s)\rangle_V ds|] \leq k$, where $k$ is independent of $l$.
- $E[\sup_{t\in I_l} | \int_0^t \langle \bar{X}^l(s), dW(s)\rangle|] \leq \frac{1}{4}E[\sup_{t\in I_l} \|X(t)\|^2] + 9E[\int_0^T \|Z(s)\|_2^2 ds]$ using the Burkholder-Davis inequality for the martingale $\int_0^t \langle \bar{X}^l(s), Z(s)dW(s)\rangle$ (which is well defined as $\bar{X}^l$ is adapted to $\mathcal{F}_t$).
- $E[\sum_{j=0}^{k_l} \| \int_{t_j^l}^{t_{j+1}^l} Z(s)dW(s)\|^2] = E[\int_0^t \|Z(s)\|_2^2 ds] \leq E[\int_0^T \|Z(s)\|_2^2 ds]$ by the Itô -isometry.

Putting this together gives

$$E[\sup_{t\in I_l} \|X(t)\|^2] \leq k + \frac{1}{4}E[\sup_{t\in I_l} \|X(t)\|^2] + 10E[\int_0^T \|Z(s)\|_2^2 ds]$$

so that

$$E[\sup_{t\in I_l} \|X(t)\|^2] \leq \tilde{k}$$

where $\tilde{k}$ is a number independent of $l$. Letting $l \to \infty$, and using the monotone convergence theorem we have

$$E[\sup_{t\in I} \|X(t)\|^2] \leq \tilde{k},$$

where $I := \bigcup_{l\in\mathbb{N}} I_l$.

Since $I$ is dense in $[0,T]$ it can also be shown that

$$\sup_{t\in I} \|X(t)\|^2 = \sup_{t\in[0,T]} \|X(t)\|^2,$$

which gives the first result.

The next step is to show that the Itô-formula holds for all $t \in I$. By letting $l \to \infty$ in (18) and (19), it holds that

- $$\int_0^t {}_{V^*}\langle Y(s), \tilde{X}^l(s)\rangle_V ds \to \int_0^t {}_{V^*}\langle Y(s), X(s)\rangle_V ds$$

- $$\sup_{t\in[0,T]} \int_0^t \langle \bar{X}^l(s), Z(s)dW(s)\rangle \to \sup_{t\in[0,T]} \int_0^t \langle X(s), Z(s)dW(s)\rangle$$

- $$\sum_{j=1}^{k_l} \|X(t_{j+1}^l) - X(t_j^l) - \int_{t_j^l}^{t_{j+1}^l} Z(s)dW(s)\|^2 \to 0,$$

where the convergence is taken in probability. There then exist a subsequence, $\{l_k\}$, and a set $\Omega' \in \mathcal{F}$ with $P(\Omega') = 1$, such that the convergence is pointwise on $\Omega'$. This gives the desired result,

$$\|X(t)\|^2 = \|X_0\|^2 + \int_0^t 2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 ds + \int_0^t 2\langle X(s), Z(s)dW(s)\rangle$$

for every $t \in I$.

To see that the formula holds even for $t \notin I$, choose a sequence $\{t_j\} \subset I$ such that $t_j < t$ and $t_j \to t$. Then using the established formula for pairs of the sequence $\{X(t_j)\}$, it follows that this is a Cauchy sequence and hence has a limit. Since $s \mapsto X(s)$ is continuously $V^*$-valued in the norm topology, it is also weak-continuously $V^*$-valued. Since $\mathcal{H}^* \subset V^*$ the map $s \mapsto X(s)$ is weak-continuously $\mathcal{H}$-valued. Then the weak and strong limit must coincide, so that actually

$$\|X(t)\|^2 = \lim_{l\to\infty} \|X(t_j^l)\|^2$$

$$= \lim_{l\to\infty} \|X_0\|^2 + \int_0^{t_j^l} 2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 ds + \int_0^{t^j} 2\langle X(s), Z(s)dW(s)\rangle$$

$$= \|X_0\|^2 + \int_0^t 2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 ds + \int_0^t 2\langle X(s), Z(s)dW(s)\rangle$$

as the functions $t \mapsto \int_0^t 2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 ds$ and $t \mapsto \int_0^t 2\langle X(s), Z(s)dW(s)\rangle$ both are continuous.

This gives the Itô-formula for every $t \in [0,T]$, and by a similar argument as above, it also follows that $X$ is continuously $\mathcal{H}$-valued. □

**Corollary 3.2.** *With $X, Y, Z$ and $X_0$ as in Theorem 3.1 , it holds that*

$$E[\|X(t)\|^2] = E\left[\|X_0\|^2 + \int_0^t 2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 ds\right]$$

### 3.2 Mild Solutions of SPDEs

Let $\{S(t)\}_{t\in[0,T]}$ be a strongly continuous semi-group on $\mathcal{H}$ with generator

$$A : \mathcal{D}(A) \to \mathcal{H}.$$

Consider the maps

$$a : [0,T] \times \mathcal{H} \to \mathcal{H}$$

$$b : [0,T] \times \mathcal{H} \to L_2(U,\mathcal{H}).$$

and the equation

$$X(t) = X_0 + \int_0^t AX(s) + a(s, X(s))ds + \int_0^t b(s, X(s))dW(s) \tag{20}$$

for some $\mathcal{F}_0$-measurable random variable $X_0$.

**Definition 3.3.** *A mild solution of (20) is an $\mathcal{F}_t$-adapted, $\mathcal{H}$-valued process, $X$, satisfying*

$$X(t) = S(t)X_0 + \int_0^t S(t-s)a(s, X(s))ds + \int_0^t S(t-s)b(s, X(s))dW(s). \tag{21}$$

**Theorem 3.4.** *Assume that $a$ and $b$ satisfies*

$$\|a(t,x) - a(t,y)\| + \|b(t,x) - b(t,y)\|_2 \leq C\|x-y\| \tag{22}$$

*for some constant $C$, and every $x, y \in \mathcal{H}$, $t \in [0,T]$. Also assume that $X_0 \in L^p(\Omega, \mathcal{F}_0, P; \mathcal{H})$, $a(\cdot, 0) \in L^p([0,T]; \mathcal{H})$ and $b(\cdot, 0) \in L^p([0,T]; L_2(U,\mathcal{H}))$ for $p \geq 2$. Then there exists a unique mild solution, $X$ to (20) such that*

$$\sup_{t\in[0,T]} (E[\|X(t)\|^p])^{1/p} < \infty.$$

To prove the theorem, a result on stochastic convolution is needed. Notice that the process $t \mapsto \int_0^t S(t-s)b(X(s))dW(s)$ is not in general a martingale, so that Doob's martingale inequality needs extension. The proof can be found in [CT06].

**Lemma 3.5 (Stochastic convolution).** *Let $p > 2$. For a process $\phi \in M^p([0,T]; L_2(U,\mathcal{H}))$ and a strongly continuous semi-group, $S(t)$, there exists a constant $C_0 > 0$ such that*

$$E[\sup_{t\in[0,T]} \| \int_0^t S(t-s)\phi(s)dW(s)\|^p] \leq C_0 E[\int_0^T \|\phi(s)\|_2^p ds] \tag{23}$$

*Proof of theorem 3.4.* The proof will be done by a fixed-point argument. Define the space $V := \{X \in L^\infty([0,T]; L^p(\Omega, \mathcal{F}, P; \mathcal{H})) : X \text{ is adapted to } \mathcal{F}_t\}$ and the mapping $G : V \to V$ given by

$$G(X)(t) = S(t)X_0 + \int_0^t S(t-s)a(s, X(s))ds + \int_0^t S(t-s)b(s, X(s))dW(s).$$

To see that $G$ is well-defined, let $M_T := \sup_{t\in[0,T]} \|S(t)\|$ which is finite by Lemma 1.28. Then

$$E[\|G(X(t))\|^p] \leq 3^{p-1}(E[\|S(t)X_0\|^p] + E[\| \int_0^t S(t-s)a(s, X(s))ds\|^p]$$

$$+E[\| \int_0^t S(t-s)b(s, X(s))dW(s)\|^p]).$$

The first summand above can be bounded by $M_T^p E[\|X_0\|^p]$. Inserting $x = X(s)$ and $y = 0$ into the Lipschitz-condition (22), it follows that

$$\|a(s, X(s))\|^p \leq 2^{p-1}\left(\|a(s,0)\|^p + C^p\|X(s)\|^p\right).$$

Using this, and the fact that $t \mapsto t^p$ is convex on $\mathbb{R}_+$, it follows that the second term can be bounded by

$$E[\int_0^t \|S(t-s)\|^p \|a(s, X(s))\|^p ds] \leq M_T^p 2^{p-1}\left(\int_0^t \|a(s,0)\|^p + C^p E[\|X(s)\|^p]ds\right)$$

$$\leq M_T^p 2^{p-1}\left(\int_0^T \|a(s,0)\|^p ds + C^p T \sup_{s\in[0,T]} E[\|X(s)\|^p]\right).$$

For the last summand, consider first at the case $p > 2$. By the stochastic convolution property (23) we get

$$E[\| \int_0^t S(t-s)b(s, X(s))dW(s)\|^p] \leq C_0 E[\int_0^t \|b(s, X(s))\|_2^p ds\|]$$

$$\leq C_0 E[2\int_0^t M_T\|b(s,0)\|_2^2 + C^2\|X(s)\|^2 ds]^{p/2}$$

$$\leq 2C_0 M_T \int_0^T \|b(s,0)\|_2^2 ds + C_0 C^2 T^{(p-2)/p} T \sup_{s\in[0,T]} E[\|X(s)\|^p],$$

similarly using the Lipschitz-condition. When $p = 2$, a similar bound is constructed using the Itô-isometry.
Putting this together gives

$$E[\|G(X)(t)\|^p] \leq C_1 + C_2 \sup_{s\in[0,T]} E[\|X(s)\|^p]$$

for suitable constants $C_1$ and $C_2$. The left hand side is independent of t, and as $X \in V$ by hypothesis, it follows that also $G(X) \in V$.

Now let $X, Y \in V$. Then

$$E[\|G(X)(t)-G(Y)(t)\|^p] \leq 2^{p-1}\left(E[\|\int_0^t S(t-s)(a(s,X(s))-a(s,Y(s)))ds\|^p]\right.$$

$$\left. + E[\|\int_0^t S(t-s)(b(s,X(s))-b(s,Y(s)))dW(s)\|^p]\right).$$

Using the Lipschitz condition on $a$, the first summand is dominated by

$$E[\int_0^t M_T^p \|a(s,X(s))-a(s,Y(s))\|^p ds] \leq M_T^p C^p E[\int_0^t \|X(s)-Y(s)\|^p ds]$$

$$\leq M_T^p C^p T \|X-Y\|_V^p.$$

Assume $p > 2$. Again, by the stochastic convolution property (23)

$$E[\|\int_0^t S(t-s)(b(s,X(s))-b(s,Y(s)))dW(s)\|^p]$$

$$\leq C_0 E[\int_0^T \|b(s,X(s))-b(s,Y(s))\|_2^p ds]$$

$$\leq C_0 E[C^p \int_0^t \|X(s)-Y(s)\|^p ds] \leq C_0 C^p T^{(p-2)/p} T \|X-Y\|_V^p.$$

For $p = 2$ the second summand is equal to

$$E[\int_0^t \|S(t-s)b(s,X(s))-S(t-s)b(s,Y(s))\|_2^2 ds] \leq C^2 T M_T \|X-Y\|_V^2.$$

Putting this together gives that for every $p \geq 2$

$$\|G(X)-G(Y)\|_V \leq \left(2^{p-1} M_T^p C^p T + 2^{p-1} C^p T^{2(p-1)/p} \max\{C_0, M_T\}\right)^{1/p} \|X-Y\|_V.$$

Let us for a moment choose $T$ such that

$$2^{p-1} M_T^p C^p T + 2^{p-1} C^p T^{2(p-1)/p} \max\{C_0, M_T\} < 1.$$

Then $G$ is a contraction mapping, so there exists a fixed point in $V$, i.e. there exists $X \in V$ such that $G(X) = X$. This is the mild solution of the SPDE. For a general $T > 0$, use the standard technique of dividing $[0,T]$ into $[0,\tilde{T}], [\tilde{T}, 2\tilde{T}], \ldots$ where $\tilde{T}$ is chosen such that $G$ is a contraction. Then, a solution can be obtained on every small interval and a solution on $[0,T]$ is obtained by gluing together these solutions.

□

### 3.3 Variational Solutions of Linear SPDE

In this section $V$ will also be assumed to be a Hilbert space. The map

$$A : V \to V^*$$

will be a bounded linear operator, satisfying

$$2_{V^*}\langle Au, u\rangle_V \le \lambda \|u\|^2 - \alpha \|u\|^2. \tag{24}$$

Consider the maps

$$b : [0, T] \times \mathcal{H} \times \Omega \to \mathcal{H}$$

$$\sigma : [0, T] \times \mathcal{H} \times \Omega \to L_2(U, \mathcal{H})$$

The main purpose of this section is to show existence and uniqueness of a solution of

$$X(t) = h + \int_0^t A(X(s)) + b(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s)$$

where $h \in \mathcal{H}$. This will be done in several steps.

**Lemma 3.6.** *Assume that $b \in M^2([0, T]; V)$ and $\sigma \in M^2([0, T]; L_2(U, V))$. Then there exists a unique solution to*

$$X(t) = h + \int_0^t A(X(s)) + b(s)ds + \int_0^t \sigma(s)dW(s) \tag{25}$$

*such that $X \in M^2([0, T]; V)$ and is continuously $\mathcal{H}$-valued.*

*Proof.* Define the adapted $V$-valued process $Y(t) = \int_0^t b(s)ds + \int_0^t \sigma(s)dW(s)$. Then, using the Itô isometry,

$$E[\|Y(t)\|_V^2] \le 2E[\int_0^T \|b(s)\|_V^2 ds] + 2E[\int_0^T \|\sigma(s)\|_{L_2(U,V)}^2 ds], \tag{26}$$

and the right-hand side is independent of $t$. Now, $AY \in M^2([0, T]; V^*)$. Indeed,

$$E[\int_0^T \|AY(s)\|_{V^*}^2 ds] \le \|A\|^2 \int_0^T E[\|Y(s)\|_V^2 ds] \le \|A\|^2 T \sup_{t\in[0,T]} E[\|Y(t)\|_V^2]. \tag{27}$$

Now choose $\Omega_0 \subset \Omega$ such that $P(\Omega_0) = 1$ and $AY \in L^2([0, T]; V^*)$ on $\Omega_0$. Then, by the deterministic equation (2.7), for every $\omega \in \Omega_0$ there exists a solution to

$$dZ(t, \omega) = (AZ(t, \omega) + AY(t, \omega))\, dt$$

$$Z(0, \omega) = h.$$

For completeness, define $Z(t,\omega) = 0$ on $[0,T] \times \Omega_0^c$. To see that $Z \in M^2([0,T];V)$, look at

$$E[\|Z(t)\|_V^2] \le 3\left(\|h\|_V^2 + \|A\|^2 E[\int_0^t \|Z(s)\|_V^2 ds] + E[\int_0^T \|AY(s)\|_{V^*}^2]]\right),$$

which by Gronwalls lemma is bounded by

$$3\left(\|h\|_V^2 + E[\int_0^T \|AY(s)\|_{V^*}^2 ds]\right) e^{\|A\|^2 T}$$

independently of $t$. Since $Y$ is $\mathcal{F}_t$-adapted, so is $Z$. Now define the process

$$X(t) := Z(t) + Y(t)$$

Then $X$ solves (25):

$$\int_0^t AX(s)ds = \int_0^t AZ(s) + AY(s)ds = Z(t) - h$$

$$= X(t) - \int_0^t b(s)ds - \int_0^t \sigma(s)dW(s) - h.$$

□

**Lemma 3.7.** *Assume that $b \in M^2([0,T];\mathcal{H})$ and $\sigma \in M^2([0,T];L_2(U,\mathcal{H}))$. Then there exists a unique solution to*

$$X(t) = h + \int_0^t A(X(s)) + b(s)ds + \int_0^t \sigma(s)dW(s)$$

*such that $X \in M^2([0,T];V)$ and is continuously $\mathcal{H}$-valued.*

*Proof.* Choose sequences $\{b_n\}$ and $\{\sigma_n\}$ in $M^2([0,T];V)$ and $M^2([0,T];L_2(U,V))$ respectively, such that

$$b_n \to b$$

and

$$\sigma_n \to \sigma$$

in $M^2([0,T];\mathcal{H})$ and $M^2([0,T];L_2(U,\mathcal{H}))$ respectively. By the previous lemma, for every $n$, there exists a solution to

$$X^n(t) = h + \int_0^t A(X^n(s)) + b_n(s)ds + \int_0^t \sigma_n(s)dW(s)$$

By Itô's formula and the condition on $A$,

$$E[\|X^n(t) - X^m(t)\|^2] = E[\int_0^t 2_{V^*}\langle A(X^n(s) - X^m(s)), X^n(s) - X^m(s)\rangle_V$$

$$+2\langle X^n(s)-X^m(s), b_n(s)-b_m(s)\rangle + \|\sigma_n(s)-\sigma_m(s)\|_2^2 ds]$$
$$\leq E[\int_0^t \lambda\|X^n(s)-X^m(s)\|^2 - \alpha\|X^n(s)-X^m(s)\|_V^2 + 2\|X^n(s)-X^m(s)\|\|b_n(s)-b_m(s)\|$$
$$+\|\sigma_n(s)-\sigma_m(s)\|_2^2 ds] \leq E[\int_0^t (\lambda+1)\|X^n(s)-X^m(s)\|^2 ds]$$
$$+p(n,m)+q(n,m)-\alpha E[\int_0^t \|X^n(s)-X^m(s)\|_V^2 ds],$$

where

$$p(n,m) := E[\int_0^T \|b_n(s)-b_m(s)\|^2 ds]$$

and

$$q(n,m) := E[\int_0^T \|\sigma_n(s)-\sigma_m(s)\|_2^2 ds],$$

which both converge to zero as $m,n \to \infty$ by the choice of $b_n$ and $\sigma_n$. It then follows by Gronwall's inequality that

$$\sup_{t\in[0,T]} E[\|X^n(t)-X^m(t)\|^2] \leq e^{(1+\lambda)T}(p(n,m)+q(n,m)) \to 0$$

as $n,m \to \infty$. Also,

$$E[\int_0^T \|X^n(s)-X^m(s)\|_V^2 ds] \leq \frac{1}{\alpha} T \sup_{t\in[0,T]} E[\|X^n(t)-X^m(t)\|^2] \to 0$$

so that $X^n$ is a Cauchy sequence in $M^2([0,T];V)$ which converges to some $X \in M^2([0,T];V)$. As $A$ is continuous, it follows that $AX^n \to AX$ in $M^2([0,T];V^*)$ and therefore

$$X(t) = h + \int_0^t A(X(s)) + b(s)ds + \int_0^t \sigma(s)dW(s)$$

as a strong limit in $M^2([0,T];V)$. Now by the statement in the Itô -formula (3.1), $X$ is also continuously $\mathcal{H}$-valued. Uniqueness follows directly from the same argument as in Theorem 2.7. $\square$

**Theorem 3.8.** *Assume*

$$b : [0,T]\times\mathcal{H}\times\Omega \to \mathcal{H}$$
$$\sigma : [0,T]\times\mathcal{H}\times\Omega \to L_2(U,\mathcal{H})$$

*are both adapted maps, and on $\Omega$ it holds that*

$$\|b(t,x)-b(t,y)\| + \|\sigma(t,x)-\sigma(t,y)\|_2 \leq C\|x-y\|,$$

*where $C$ is independent of $t$. Then there exists a unique $X \in M^2([0,T];V)$ which is continuously $\mathcal{H}$-valued and*

$$X(t) = h + \int_0^t A(X(s)) + b(s,X(s))ds + \int_0^t \sigma(s,X(s))dW(s).$$

*Proof.* First define

$$X^0(t) = h$$

and inductively,

$$X^{n+1}(t) = h + \int_0^t AX^{n+1}(s) + b(s, X^n(s))ds + \int_0^t \sigma(s, X^n(s))dW(s)$$

for every $n = 1, 2, \ldots$. By the previous lemma, such a solution exists uniquely in $M^2([0,T]; V)$. By Itô's formula, similarly as in the proof of the previous lemma,

$$E[\|X^{n+1}(t) - X^n(t)\|^2] + \alpha E[\int_0^t \|X^{n+1}(s) - X^n(s)\|_V^2 ds]$$

$$\leq \int_0^t (1+\lambda) E[\|X^{n+1}(s) - X^n(s)\|^2] ds$$

$$+ \int_0^t E[\|b(s, X^n(s)) - b(s, X^{n-1}(s))\|^2 + \|\sigma(s, X^n(s)) - \sigma(s, X^{n-1}(s))\|_2^2] ds$$

$$\leq \int_0^t (1+\lambda) E[\|X^{n+1}(s) - X^n(s)\|^2] ds + \int_0^t C E[\|X^n(s) - X^{n-1}(s)\|^2] ds.$$

Define the functions

$$f_n(t) := E[\|X^{n+1}(t) - X^n(t)\|^2]$$

and

$$g_n(t) := E[\int_0^t \|X^{n+1}(s) - X^n(s)\|_V^2 ds],$$

so that

$$f_n(t) + \alpha g_n(t) \leq (1+\lambda) \int_0^t f_n(s) + C f_{n-1}(s) ds$$

for every n. By Gronwall's inequality, it follows that

$$f_n(t) \leq e^{(1+\lambda)t} \int_0^t C f_{n-1}(s) ds,$$

and by induction

$$f_n(t) \leq \frac{(\tilde{C}t)^n}{n!} \leq \frac{(\tilde{C}T)^n}{n!}$$

for a suitable constant $\tilde{C}$, and hence

$$\sup_{t \in [0,T]} f_n(t) \to 0$$

as $n \to \infty$. Then

$$g_n(T) \leq \alpha^{-1} \int_0^T (1+\lambda) f_n(s) + C f_{n-1}(s) ds$$

$$\leq \alpha^{-1}T((1+\lambda)\sup_t f_n(t) + C\sup_t f_{n-1}(t)) \to 0$$

as $n \to \infty$. Now by the same argument as in the previous proof, and by the continuity of $A, b$ and $\sigma$, the result follows.

□

The above proofs are merely algebraic manipulations by adjoining relations on the coefficients, and the real work is done in the deterministic case 2.7. The following example is of the same nature.

**Example 3.9 (Ornstein Uhlenbeck equation).** *Let $B \in L_2(U, L^2(\Lambda))$ be constant. Then there exists a variational solution to*

$$X(t,x) = X_0(x) + \int_0^t \Delta X(s,x)ds + BW(t)$$

*on the Gelfand triple $H_0^1(\Lambda) \subset L^2(\Lambda) \subset H^{-1}(\Lambda)$. This process is called an infinite-dimensional Ornstein-Uhlenbeck equation.*

*Proof.* From the proof of Example 2.8 it is clear that $\Delta$ satisfies the condition to guarantee a solution. Also, since $B$ is constant, the result follows. □

### 3.4 Variational Solutions of non-linear SPDE

This section will be devoted to discuss existence and uniqueness of equations of the form

$$X(t) = X_0 + \int_0^t A(s, X(s))ds + \int_0^t B(s, X(s))dW(s),$$

where $A$ and $B$ may be non-linear operators. The setup is as follows:

- $A : [0,T] \times V \times \Omega \to V^*$
  $B : [0,T] \times V \times \Omega \to L_2(U, \mathcal{H})$
  are both adapted maps.
- $X_0 \in L^2(\Omega; V)$ is $\mathcal{F}_0$-measurable.

Given $(A, B, X_0)$ as above, a solution is a continuous, $\mathcal{F}_t$-adapted process

$$X : [0,T] \times \Omega \to \mathcal{H}$$

such that $X \in M^\alpha([0,T]; V) \cap M^2([0,T]; \mathcal{H})$ for some $\alpha > 1$, and satisfies

$$X(t) = X_0 + \int_0^t A(s, X(s))ds + \int_0^t B(s, X(s))dW(s).$$

Notice how the definition of a solution depends on $\alpha$. This constant will be chosen a posteriori in order to fit the solution to the conditions needed on $A$ and $B$ as in the following theorem:

**Theorem 3.10.** *Let $X_0 \in L^2(\Omega; \mathcal{H})$ be $\mathcal{F}_0$-measurable. Let $A$ and $B$ be as above, and in addition, assume that*

1. *For all $u, v, x \in V$ and $t \in [0,T]$ the map*
$$\begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ \lambda & \mapsto & {}_{V^*}\langle A(t, u + \lambda v), x\rangle_V \end{array}$$
*is continuous on $\Omega$.*

2. *There exists $c \in \mathbb{R}$ such that for any $u, v \in V$ and $t \in [0,T]$,*
$$2_{V^*}\langle A(t,u) - A(t,v), u - v\rangle_V + \|B(t,u) - B(t,v)\|_2^2 \leq c\|u - v\|^2$$
*holds on $\Omega$.*

3. *There exists scalars $c_1 \in \mathbb{R}$ and $c_2 \in \mathbb{R}_+$ and an $\mathcal{F}_t$-adapted process $f \in L^1([0,T] \times \Omega)$ such that for every $v \in V$ and $t \in [0,T]$,*
$$2_{V^*}\langle A(t,v), v\rangle_V + \|B(t,v)\|_2^2 \leq c_1\|v\|^2 - c_2\|v\|_V^\alpha + f(t)$$

4. *There exists a scalar $c_3 \in \mathbb{R}_+$ and an $\mathcal{F}_t$-adapted process $g \in L^{\alpha/(1-\alpha)}([0,T] \times \Omega)$ such that for every $v \in V$ and $t \in [0,T]$,*
$$\|A(t,v)\|_{V^*} \leq c_3\|v\|_V^{\alpha-1} + g(t).$$

*Then there exists a unique solution as described above.*

*Proof.* As in the deterministic case, the trick is to make finite-dimensional approximations of the solution, and then extract a subsequence that converges in a weak topology. The proof will be closed by showing that this weak limit is in fact a solution to the problem. As before, define the spaces $\mathcal{H}_n = \text{span}\{e_1, \dots e_n\}$ for every $n \in \mathbb{N}$ where $\{e_n\}_{n\in\mathbb{N}}$ is an orthonormal sequence in $\mathcal{H}$ and $\text{span}\{e_n\}_{n\in\mathbb{N}}$ is dense in $V$. Let also $P_n$ be as before. Define $W^n(t) := \sum_{k=1}^n B^k(t) f_k$. For an orthonormal basis $\{f_n\}_{n\in\mathbb{N}}$, $U_n = \text{span}\{f_1, \dots f_n\}$ and $\tilde{P}_n$ the orthogonal projection from $U$ to $U_n$ we have $W^n(t) = \tilde{P}_n W(t)$.

Look now at equations of the form
$$dX^n(t) = P_n A(t, X^n(t))dt + P_n B(t, X^n(t))dW^n(t)$$
$$X^n(0) = P_n X_0.$$
Where the maps can be considered as
$$P_n A : [0,T] \times \mathcal{H}_n \times \Omega \to \mathcal{H}_n$$
$$P_n B : [0,T] \times \mathcal{H}_n \times \Omega \to L_2(U_n, \mathcal{H}_n).$$

Now since all the spaces above are finite-dimensional, $\mathcal{H}_n \simeq \mathbb{R}^n$, $U_n \simeq \mathbb{R}^n$ and so $L_2(U_n, \mathcal{H}_n) \simeq \mathbb{R}^{n\times n}$, the space of all $n \times n$-matrices. Making these identifications we arrive at an equation on the form

$$d\tilde{X}^n(t) = \tilde{A}(t, \tilde{X}^n(t))dt + \tilde{B}(t, \tilde{X}^n(t))dW^n(t)$$
$$\tilde{X}^n(0) = \widetilde{P_n X_0}$$

where

$$\tilde{A} : [0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^n$$
$$\tilde{B} : [0,T] \times \mathbb{R}^n \times \Omega \to \mathbb{R}^{n\times n}$$

Now this is an ordinary stochastic differential equation, and by [PR07] there exists a solution to this equation for every $n$. Here we have used the conditions 1 to 4. Let now $X^n$ be the stochastic process on $\mathcal{H}_n$ via the natural embedding of $\tilde{X}^n$ from $\mathbb{R}^n$ to $\mathcal{H}_n$. Then this process satisfies

$$dX^n(t) = P_n A(t, X^n(t))dt + P_n B(t, X^n(t))dW^n(t)$$

$$X^n(0) = P_n X_0$$

Using the finite-dimensional Itô-formula, the (deterministic) product rule on $e^{-c_1 t}E[\|X^n(t)\|^2]$ and assumption 3 it is possible to show that

$$\|X^n\|_{M^\alpha([0,T];V)} + \sup_{t\in[0,T]} E[\|X^n(t)\|^2] \leq C$$

for some constant $C$ which is independent of $n$. By assumption 3 and 4 on $A$ and $B$, it then follows that also

$$\|A(\cdot, X^n)\|_{M^{\alpha/(\alpha-1)}([0,T];V^*)} + \|B(\cdot, X^n)\|^2_{M^2([0,T];L_2(U,\mathcal{H}))} \leq \bar{C}$$

for another constant $\bar{C}$, still independent of $n$. Now, as these sequences are bounded, there exists elements

$$X \in M^\alpha([0,T];V) \cap M^2([0,T];\mathcal{H})$$

$$Y \in M^{\alpha/(\alpha-1)}([0,T];V^*)$$

$$Z \in M^2([0,T];L_2(U,\mathcal{H}))$$

such that

- $X^n \to X$ weakly in $M^\alpha([0,T];V)$ and $M^2([0,T];\mathcal{H})$
- $A(\cdot, X^n) \to Y$ weakly in $M^{\alpha/(\alpha-1)}([0,T];V^*)$
- $P_n B(\cdot, X^n) \to Z$ weakly in $Z \in M^2([0,T];L_2(U,\mathcal{H}))$

as $n \to \infty$, for some subsequence (still denoted by $n$). It follows that

$$\sup_{t\in[0,T]} \int_0^t P_n B(s, X^n(s)) dW^n(s) \to \sup_{t\in[0,T]} \int_0^t Z(s) dW(s)$$

weakly in $L^2(\Omega; \mathcal{H})$, and so

$$X(t) = X_0 + \int_0^t Y(s)ds + \int_0^t Z(s)dW(s)$$

as a weak limit in $M^\alpha([0,T]; V) \cap M^2([0,T]; \mathcal{H})$.
The (infinite-dimensional version of the) Itô-formula and the (deterministic) product rule gives

$$e^{-ct} E[\|X(t)\|^2] =$$

$$E[\|X_0\|^2] + \int_0^t e^{-cs} E\left[\left(2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 - c\|X(s)\|^2\right)\right] ds.$$

As $X(t)$ is a weak limit, it holds that

$$\|X\|_{M^\alpha([0,T];V)} \leq \liminf_{n\to\infty} \|X^n(t)\|_{M^\alpha([0,T];V)}$$

and for a $h \in L^1([0,T]; \mathbb{R}_+)$, we still get that $hX^n \to hX$ weakly, and so,

$$\|hX\|_{M^\alpha([0,T];V)} \leq \liminf_{n\to\infty} \|hX^n(t)\|_{M^\alpha([0,T];V)}.$$

The same argument also gives that

$$\|h(t)X(t)\|^2 \leq \liminf_{n\to\infty} \|h(t)X^n(t)\|^2.$$

Applying this to $h(t) = e^{-ct}$ gives

$$\int_0^t e^{-cs} E\left[\left(2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 - c\|X(s)\|^2\right)\right] ds$$

$$\leq \int_0^t e^{-cs} E\left[\left(2_{V^*}\langle A(s, X^n(s)), X(s)\rangle_V + \|B(s, X^n(s))\|_2^2 - c\|X(s)\|^2\right)\right] ds$$

Let now $\phi \in M^\alpha([0,T]; V) \cap M^2([0,T]; \mathcal{H})$ be arbitrary. By the trivial equalities (for notational convenience the time-variable is removed from the equations)

- ${}_{V^*}\langle A(X^n), X^n\rangle_V = {}_{V^*}\langle A(X^n) - A(\phi), X^n - \phi\rangle_V$
  $+{}_{V^*}\langle A(X^n) - A(\phi), \phi\rangle_V + {}_{V^*}\langle A(\phi), X^n\rangle_V$
- $\|B(X^k)\|_2^2 = \|B(X^n) - B(\phi)\|_2^2 + 2\langle B(X^n), B(\phi)\rangle_2$
  $-\|B(\phi)\|_2^2$
- $\|X^n\|^2 = \|X^n - \phi\|^2 + 2\langle X^n, \phi\rangle - \|\phi\|^2$

and

$$\langle A(X^n) - A(\phi), X^n - \phi\rangle + \|B(X^k) - B(\phi)\|_2^2 \leq c\|X^k - \phi\|$$

(from the second assumption), all inserted into the above inequality gives

$$E\left[\int_0^t e^{-cs}\left(2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 - c\|X(s)\|^2\right) ds\right]$$

$$\leq E\left[\int_0^t e^{-cs}\left(2_{V^*}\langle A(s,\phi(s)), X^n(s)\rangle_V + 2_{V^*}\langle A(s,X^n(s)) - A(s,\phi(s)), \phi(s)\rangle_V\right.\right.$$

$$\left.\left. -\|B(s,\phi(s))\|_2^2 + 2\langle B(s,X^n(s)), B(s,\phi(s))\rangle_2 - 2c\langle X^n(s), \phi(s)\rangle + c\|\phi(s)\|^2\right) ds\right]$$

Letting $n \to \infty$ then gives the following

$$E\left[\int_0^t e^{-cs}\left(2_{V^*}\langle Y(s), X(s)\rangle_V + \|Z(s)\|_2^2 - c\|X(s)\|^2\right) ds\right]$$

$$\leq E\left[\int_0^t e^{-cs}\left(2_{V^*}\langle A(s,\phi(s)), X(s)\rangle_V + 2_{V^*}\langle Y(s)) - A(s,\phi(s)), \phi(s)\rangle_V\right.\right.$$

$$\left.\left. -\|B(s,\phi(s))\|_2^2 + 2\langle Z(s)), B(s,\phi(s))\rangle_2 - 2c\langle X(s), \phi(s)\rangle + c\|\phi(s)\|^2\right) ds\right]$$

Since $\phi \in M^\alpha([0,T];V) \cap M^2([0,T];\mathcal{H})$ was arbitrary, it then follows that

- $Y(t) = A(t, X(t))\ dt \times P$-a.s., and
- $Z(t) = B(t, X(t))\ dt \times P$-a.s.

This proves existsence of the solution. To show uniqueness, assume that $X$ and $Y$ are two solutions. Using the Itô-formula to the difference process $X - Y$ and taking expectation, it follows that

$$E[\|X(t) - Y(t)\|^2] \leq E[\int_0^t 2_{V^*}\langle A(s,X(s)) - A(s,Y(s)), X(s) - Y(s)\rangle_V$$

$$+\|B(s,X(s)) - B(s,Y(s))\|_2^2 ds] \leq c\int_0^t E[\|X(s) - Y(s)\|^2]ds,$$

where the last inequality comes from assumption 2. From Gronwall's inequality, it then follows that

$$E[\|X(t) - Y(t)\|^2] = 0$$

so that the solution is P-a.s. unique, for every $t \in [0,T]$.

□

**Example 3.11.** *Consider the equation*

$$dX(t) = div\left(|\nabla X(t)|^{p-2}\nabla X(t)\right) dt + BdW(t).$$

*This equation has a unique solution on the Gelfand-triple*

$$W_0^{1,p}(\Lambda) \subset L^2(\Lambda) \subset (W_0^{1,p}(\Lambda))^*$$

*where $\Lambda$ is a bounded open subset of $\mathbb{R}^n$ and divergence is taken in the sense of distribution, i.e. for a vector field $F : \mathbb{R}^n \to \mathbb{R}^n$*

$$\langle div(F), \varphi\rangle = -\int_\Lambda \langle F, \nabla\varphi\rangle d\lambda$$

*for $\varphi \in W_0^{1,p}(\Lambda)$.*

*Proof.* First, let $f, g \in L^p(\Lambda)$. It then follows that $f^{p-1}g \in L^1(\Lambda)$. Indeed, by Hölder‘s inequality

$$\int_\Lambda |f|^{p-1}|g|d\lambda \leq \left(\int_\Lambda |g|^p d\lambda\right)^{\frac{p-1}{p}} \|f\|_p = \|g\|_p^{p-1}\|f\|_p. \tag{28}$$

Define now the operator $A : W_0^{1,p}(\Lambda) \to (W_0^{1,p}(\Lambda))^*$ by

$$\langle Au, v\rangle = -\int_\Lambda |\nabla u|^{p-2}\langle\nabla u, \nabla v\rangle d\lambda.$$

To see that this is a well-defined map, note that by (28) with $g = |\nabla u|$ and $f = |\nabla v|$, we have

$$\int_\Lambda |\nabla u|^{p-2}\langle\nabla u, \nabla v\rangle d\lambda \leq \int_\Lambda |\nabla u|^p|\nabla v|d\lambda$$

$$\leq \|\nabla u\|_p^{p-1}\|\nabla u\|_p \leq \|u\|_{1,p}^{p-1}\|v\|_{1,p}.$$

This also proves that

$$\|A(u)\|_{V^*} \leq \|u\|_{1,p}^{p-1}$$

This immediately gives condition 4 in the theorem, with $c_3 = 1$, $\alpha = p$ and $g = 0$. Look now at the other conditions from the theorem:

1. The first condition is equivalent to having

$$\langle A(u + n^{-1}w), v\rangle \to \langle A(u), v\rangle$$

as $n \to \infty$ for all $u, v, w \in W_0^{1,p}(\Lambda)$. To show this, look at

$$|\langle A(u+n^{-1}w) - A(u), v\rangle| \leq \int_\Lambda \big||\nabla(u + n^{-1}w)|^{p-2}\langle\nabla(u + n^{-1}w), \nabla v\rangle$$

$$-|\nabla u|^{p-2}\langle \nabla u, \nabla v\rangle\big|\, d\lambda.$$

Clearly,

$$\big||\nabla(u+n^{-1}w)|^{p-2}\langle \nabla(u+n^{-1}w), \nabla v\rangle - |\nabla u|^{p-2}\langle \nabla u, \nabla v\rangle\big| \to 0$$

as $n \to \infty$. This is dominated by

$$\begin{aligned}&|\nabla(u+n^{-1}w)|^{p-1}|\nabla v| + |\nabla u|^{p-1}|\nabla v| \\ &\leq 2^{p-2}\left(|\nabla u|^{p-1} + |\nabla w|^{p-1}\right)\nabla v + |\nabla u|^{p-1}|\nabla v|,\end{aligned}$$

and by (28), this is in $L^1(\Lambda)$. By dominated convergence theorem, it follows that

$$\int_\Lambda \big||\nabla(u+n^{-1}w)|^{p-2}\langle \nabla(u+n^{-1}w), \nabla v\rangle - |\nabla u|^{p-2}\langle \nabla u, \nabla v\rangle\big|\, d\lambda \to 0$$

as $n \to \infty$, as desired.

2. Let $u, v \in W_0^{1,p}(\Lambda)$.

$$\begin{aligned}\langle A(u) - A(v), u - v\rangle &= \langle A(u), u\rangle - \langle A(u), v\rangle - \langle A(v), u\rangle + \langle A(v), v\rangle \\ &= \int_\Lambda |\nabla u|^{p-2}\langle \nabla u, \nabla v\rangle + |\nabla v|^{p-2}\langle \nabla v, \nabla u\rangle - |\nabla u|^p - |\nabla v|^p d\lambda \\ &\leq \int_\Lambda |\nabla u|^{p-1}|\nabla v| + |\nabla v|^{p-1}|\nabla u| - |\nabla u|^p - |\nabla v|^p d\lambda \\ &= \int_\Lambda -\left(|\nabla u|^{p-1} - |\nabla v|^{p-1}|\right)\left(|\nabla u| - |\nabla v|\right) d\lambda.\end{aligned}$$

When $p > 1$, the function $t \mapsto t^{p-1}$ is increasing on $\mathbb{R}_+$, so that $(s-t)$ and $(s^{p-1} - t^{p-1})$ always has the same sign. This gives that the above integrand is always negative, so

$$\langle A(u) - A(v), u - v\rangle \leq 0$$

and condition 2 holds with $c = 0$.

3. By Poincarè's inequality (see [Eva98]) since $\Lambda$ is bounded, there exists a constant $C$ which only depends on $\Lambda$, $n$ and $p$ such that

$$\int_\Lambda |u|^p d\lambda \leq C \int_\Lambda |\nabla u|^p d\lambda$$

for all $u \in W_0^{1,p}(\Lambda)$. Then

$$\begin{aligned}\langle A(u), u\rangle &= -\int_\Lambda |\nabla u|^p d\lambda = -\left(\frac{1}{2}\int_\Lambda |\nabla u|^p d\lambda + \frac{1}{2}\int_\Lambda |\nabla u|^p d\lambda\right) \\ &\leq -\frac{1}{2C}\int_\Lambda |\nabla u|^p d\lambda - \frac{1}{2}\int_\Lambda |u|^p d\lambda \leq -\frac{1}{2}\min\{C^{-1}, 1\}\|u\|_{1,p}^p,\end{aligned}$$

so that condition 3 holds with $c_1 = 0$, $c_2 = \min\{C^{-1}, 1\}$ and $f = 0$.

Condition number 4 was proved earlier.
Hence, by the theorem, there exists a unique process

$$X \in M^2([0,T]; L^2(\Lambda)) \cap M^p([0,T]; W_0^{1,p}(\Lambda))$$

and $X$ is continuously $L^2(\Lambda)$-valued, such that

$$X(t,x) = X_0(x) + \int_0^t \operatorname{div}(|\nabla X(s,x)|^{p-2}\nabla X(s,x))ds + \int_0^t B(s)dW(s)$$

□

## 3.5 Backward SPDE

Let $B(t)$ be a one-dimensional Brownian motion on $(\Omega, \mathcal{F}, P)$ with filtration $\mathcal{F}_t$ generated by $B(t)$. A backward stochastic differential equation (BSDE) is an equation of the form

$$dY(t) = -b(t, Y(t), Z(t))dt + Z(t)dB(t)$$

$$Y(T) = \phi$$

where the terminal value of the process $Y$ is given, rather than the starting point. A solution to this equation is a pair of adapted processes $(Y, Z)$ that satisfies

$$Y(t) = \phi + \int_t^T b(s, Y(s), Z(s))ds - \int_t^T Z(s)dB(s)$$

It is well known (see e.g. [NKQ97]) that such solutions exists uniquely when $b$ is a Lipschitz function. Such equations appear in several real-life problems. In particular in finance, the problem of finding a replicating portfolio for a contingent claim can be rephrased as a BSDE. The aim of this section is to show existence and uniqueness for a small class of semi-linear backward stochastic partial differential equations (BSPDE). The setting is as follows

- A Gelfand triple $V \subset \mathcal{H} \subset V^*$ where $V$ is a Hilbert space,
- $W$ is a cylindrical Brownian motion on a separable Hilbert space $U$,
- $A$ is a continuous bounded linear operator, $A : V \to V^*$, and
- $\phi \in L^2(\Omega, \mathcal{F}_T, P; \mathcal{H})$.

**Theorem 3.12.** *Assume* $2\langle Av, v\rangle \leq \lambda\|v\|^2 - \alpha\|v\|_V^2$ *for some constants* $\alpha > 0$ *and* $\lambda \geq 0$ *and every* $v \in V$. *Let*

$$b : [0,T] \times V \times L_2(U, \mathcal{H}) \times V^* \times \Omega \to \mathcal{H}$$

*be uniformly Lipschitz, i.e. for every* $y, \bar{y} \in V$, $z, \bar{z} \in L_2(U, \mathcal{H})$ *and* $w, \bar{w} \in V^*$,

$$\|b(t, y, z, w) - b(t, \bar{y}, \bar{z}, \bar{w})\| \leq C \left(\|y - \bar{y}\|_V + \|z - \bar{z}\|_2 + \|w - \bar{w}\|_{V^*}\right)$$

*for some constant* $C > 0$, *and* $b(\cdot, 0, 0, 0) \in M^2([0, T]; \mathcal{H})$. *Then there exists a unique pair* $(Y, Z) \in M^2([0, T]; V) \times M^2([0, T]; L_2(U, \mathcal{H})$ *such that*

$$Y(t) = \phi + \int_t^T AY(s) + b(s, Y(s), Z(s), AY(s))ds - \int_t^T Z(s)dW(s)$$

*for all* $t \in [0, T]$.

*Proof.* It was shown in [BØP05] that when $b(t, y, z, w) = b(t, y, z)$ is independent of $w$, there exists a unique solution to

$$Y(t) = \phi + \int_t^T AY(s) + b(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s)$$

in the same manner as described above.
Define now $(Y^0, Z^0)$ to be the unique solution to

$$Y^0(t) = \phi - \int_t^T Z^0(s)dW(s)$$

and inductively define $(Y^{n+1}, Z^{n+1})$ as the solution of

$$Y^{n+1}(t) = \phi + \int_t^T AY^{n+1}(s) + b(s, Y^{n+1}(s), Z^{n+1}(s), AY^n(s))ds$$

$$- \int_t^T Z^{n+1}(s)dW(s).$$

By Itô's formula,

$$E[\|Y^{n+1}(t) - Y^n(t)\|^2] + E[\int_t^T \|Z^{n+1}(s) - Z^n(s)\|_2^2 ds]$$

$$= E[\int_t^T 2\langle A(Y^{n+1}(s) - Y^n(s)), Y^{n+1}(s) - Y^n(s)\rangle$$

$$+2\langle b(s, Y^{n+1}(s), Z^{n+1}(s), AY^n(s)) - b(s, Y^n(s), Z^n(s), AY^{n-1}(s)), Y^{n+1}(s) - Y^n(s)\rangle ds].$$

Now by the condition on $A$, and the Cauchy-Schwartz inequality this is dominated by

$$E[\int_t^T \lambda\|Y^{n+1}(s) - Y^n(s))\|^2 - \alpha\|Y^{n+1}(s) - Y^n(s)\|_V^2$$

$+2\|b(s,Y^{n+1}(s),Z^{n+1}(s),AY^n(s))-b(s,Y^n(s),Z^n(s),AY^{n-1}(s))\|\|Y^{n+1}(s)-Y^n(s)\|ds].$

By the Lipschitz-condition, the second integrand is dominated by

$$2C\left(\|Y^{n+1}(s)-Y^n(s)\|_V+\|Z^{n+1}(s)-Z^n(s)\|_2\right.$$

$$\left.+\|AY^n(s)-AY^{n-1}(s)\|_{V^*}\right)\|Y^{n+1}(s)-Y^n(s)\|.$$

Now using the inequality $2ab\leq\beta a^2+\beta^{-1}b^2$ repeatedly, the above is dominated by

$$C\beta\|Y^{n+1}(s)-Y^n(s)\|_V^2+C\gamma\|Z^{n+1}(s)-Z^n(s)\|_2^2$$

$$+C\rho\|A\|^2\|Y^n(s)-Y^{n-1}(s)\|_V^2+C\left(\beta^{-1}+\gamma^{-1}+\rho^{-1}\right)\|Y^{n+1}(s)-Y^n(s)\|^2$$

for positive constants $\beta,\gamma$ and $\rho$ which will be chosen later. Putting all this together gives the following inequality

$$E[\|Y^{n+1}(t)-Y^n(t)\|^2]+E[\int_t^T\|Z^{n+1}(s)-Z^n(s)\|_2^2ds]$$

$$\leq E[\int_t^T\tilde{\lambda}\|Y^{n+1}(s)-Y^n(s)\|^2+(C\beta-\alpha)\|Y^{n+1}(s)-Y^n(s)\|_V^2$$

$$C\gamma\|Z^{n+1}(s)-Z^n(s)\|_2^2+C\rho\|A\|^2\|Y^n(s)-Y^{n-1}(s)\|_V^2ds],$$

where $\tilde{\lambda}=\tilde{\lambda}(\lambda,C,\beta,\gamma,\rho):=\lambda+C\left(\beta^{-1}+\gamma^{-1}+\rho^{-1}\right)$.
Choose now $\beta<\alpha/C$ and $\gamma<1/2C$ so that

$$E[\|Y^{n+1}(t)-Y^n(t)\|^2]+\frac{1}{2}E[\int_t^T\|Z^{n+1}(s)-Z^n(s)\|_2^2ds]$$

$$+(\alpha-C\beta)E[\int_t^T\|Y^{n+1}(s)-Y^n(s)\|_V^2ds]\leq$$

$$E[\int_t^T\tilde{\lambda}\|Y^{n+1}(s)-Y^n(s)\|^2+C\rho\|A\|^2\|Y^n(s)-Y^{n-1}(s)\|_V^2ds].$$

Define the functions

$$f_n(t)=E[\|Y^{n+1}(t)-Y^n(t)\|^2]$$

and

$$G_n(t)=E[\int_t^T\|Y^{n+1}(s)-Y^n(s)\|_V^2ds]$$

so that

$$f_n(t)+(\alpha-C\beta)G_n(t)\leq\int_t^T\tilde{\lambda}f_n(s)ds+C\rho\|A\|^2G_{n-1}(t).$$

Then by Gronwall's inequality,

$$f_n(t) \leq e^{\tilde{\lambda}t} C\rho \|A\|^2 G_{n-1}(t) \leq e^{\tilde{\lambda}T} C\rho \|A\|^2 G_{n-1}(0),$$

so that

$$(\alpha - C\beta) G_n(0) \leq T e^{\tilde{\lambda}T} C\rho \|A\|^2 G_{n-1}(0) + C\rho \|A\|^2 G_{n-1}(0).$$

Choose now $\rho < \frac{\alpha - C\beta}{2C\|A\|^2}$. Note that none of the constants involved depend on $T$. For a moment, let $T$ be such that $T e^{\tilde{\lambda}T} < \frac{1}{2}$. Then

$$G_n(0) \leq \frac{C\rho\|A\|^2}{\alpha - C\beta} G_{n-1}(0)(T e^{\tilde{\lambda}T} + 1) < \frac{1}{2} G_{n-1}(0) \left( \frac{1}{2} + 1 \right) = \frac{3}{4} G_{n-1}(0),$$

and by induction,

$$G_n(0) < \left( \frac{3}{4} \right)^n G_0(0) \to 0$$

as $n \to \infty$.

Then there exists a $Y \in M^2([0,T];V)$ such that

$$Y^n \to Y \quad \text{in } M^2([0,T];V).$$

It also follows that there exists a $Z \in M^2([0,T]; L_2(U,\mathcal{H}))$ such that

$$Z^n \to Z \quad \text{in } M^2([0,T]; L_2(U,\mathcal{H})).$$

Now by the continuity of $A$ and $b$ it follows that

$$Y(t) = \phi + \int_t^T AY(s) + b(t, Y(s), Z(s), AY(s))ds - \int_t^T Z(s)dW(s)$$

as desired.
For a general $T$, divide the interval $[0,T]$ into subintervals

$$[0, T_0], [T_0, 2T_0], \ldots, [T - T_0, T]$$

such that there exists solutions on every interval. Construct then a solution first on $[T - T_0, T]$, and solve backwards on every subinterval (this is the same technique as used in [MY07]).

To see uniqueness, assume that $(\tilde{Y}, \tilde{Z})$ is another solution to the BSPDE. Using Itô's formula and the same technique as earlier, we have that

$$E[\|Y(t) - \tilde{Y}(t)\|^2] + \frac{1}{2} E[\int_t^T \|Z(s) - \tilde{Z}(s)\|_2^2 ds]$$

$$+(\alpha - C\beta) E[\int_t^T \|Y(s) - \tilde{Y}(s)\|_V^2 ds]$$

$$\leq \tilde{\lambda} E[\int_t^T \|Y(s) - \tilde{Y}(s)\|^2] + C\rho\|A\|^2 E[\int_t^T \|Y(s) - \tilde{Y}(s)\|_V^2],$$

and with the same choice of $\rho < \frac{\alpha - C\beta}{2C\|A\|^2}$,

$$E[\|Y(t) - \tilde{Y}(t)\|^2] + \frac{1}{2} E[\int_t^T \|Z(s) - \tilde{Z}(s)\|_2^2 ds]$$

$$+\frac{1}{2} E[\int_t^T \|Y(s) - \tilde{Y}(s)\|_V^2 ds] \leq \tilde{\lambda} E[\int_t^T \|Y(s) - \tilde{Y}(s)\|^2].$$

Hence, by Gronwall's inequality it follows that $(Y, Z) = (\tilde{Y}, \tilde{Z})$ in $M^2([0,T]; V) \times M^2([0,T]; L_2(U, \mathcal{H}))$. $\square$

# 4 Applications to Interest Rates

This section will give a short presentation on how to model interest rates by means of SPDE theory. The presentation will be almost purely mathematical. Discussions on risk neutral measures and market assumptions such as no arbitrage will be left out (in fact, no arbitrage will be a mathematical condition). Instead I will try to get as quickly as possible to the equations used in modeling of the interest rates.

A *zero coupon bond* with *maturity* T, is a contract in which the holder of the contract is guaranteed \$ 1 at time T. The price of such a contract at time $t \leq T$ will be denoted by $p(t, T)$. In the following, it will be assumed that there exists a market for zero coupon bonds. We assume that $p$ is a stochastic process on some probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathcal{F}_t$ and

- The process $t \mapsto p(t, T)$ is adapted to the filtration $\mathcal{F}_t$ for each $T > 0$,
- $p(t, t) = 1$ for every $t$, and
- The map $T \mapsto p(t, T)$is P-a.s. differentiable.

With these assumptions, define the *instantaneous forward rate* as

$$f(t, T) := -\frac{\partial}{\partial T} \ln p(t, T)$$

and the *instantaneous short rate* as $r(t) := f(t, t)$.

There exists several models for short rate, e.g. the Vasicek-model. Here we let $B$ be a real-valued Brownian motion and assume that the short rate evolves according to

$$dr(t) = (b - ar(t))dt + \sigma dB(t)$$

for constants $b, a$ and $\sigma\cdot$
For the forward rate, we have the Heath-Jarrow-Morton (HJM) model. Here let $T > 0$ be fixed, and $B$ be a real-valued Brownian motion. In the HJM-model, we assume that the forward rate evolves according to

$$df(t, T) = \alpha(t, T)dt + \sigma(t, T)dB(t)$$

for processes $t \mapsto \alpha(t, T)$ and $t \mapsto \sigma(t, T)$. The HJM no-arbitrage condition states that

$$\alpha(t, T) = \sigma(t, T) \int_t^T \sigma(t, u)du,$$

so that the forward rate is entirely described by the initial forward curve $f(0, T)$ and the volatility structure $\sigma(t, T)$.

In this model, $T$ was fixed, and the Brownian motion depend on $T$. Letting now $T$ vary, we get a parametrized family of stochastic processes, and an infinite number of Brownian motions.

Define the stochastic process $X(t,x) := f(t,t+x)$ where now $x = T - t$ is the time *to* maturity. In the generalized HJM-model, $X$ is modeled by mild solutions of the equation

$$X(t,x) = X_0(x) + \int_0^t \frac{\partial X(s,x)}{\partial x} + \alpha(s,X,x)ds + \sum_{k=1}^{\infty} \int_0^t \sigma^k(s,X,x)dB^k(s) \tag{29}$$

where $\{B^k\}_{k=1}^{\infty}$ is a sequence of independent Brownian motions, and $\sigma^k$ is a sequence of processes, which can depend on $X$. The generalized HJM no-arbitrage for a function $h(x)$ now reads:

$$\alpha(t,h,x) = \sum_{k=1}^{\infty} \sigma^k(t,h,x) \int_0^x \sigma^k(t,h,u)du$$

Equation (29) will be regarded as an $\mathcal{H}$-valued equation

$$dX(t) = (AX(t) + \alpha(t,X(t)))dt + \sigma(t,X(t))dW(t) \tag{30}$$

where $A$ generates the semi-group of right translation, $\alpha$ and $\sigma$ are operators and $\mathcal{H} \subset C(\mathbb{R}_+)$ is an appropriately chosen Hilbert-space of functions. Based on the discussion above, we see that $\mathcal{H}$ should satisfy

1. The evaluation map $\delta_x : \mathcal{H} \to \mathbb{R}$ defined by $\delta_x(h) = h(x)$ is a continuous functional,
2. the integration map $I_x : \mathcal{H} \to \mathbb{R}$ defined by $I_x(h) = \int_0^x h(u)du$ is a continuous functional,
3. the semi-group of left translation, $S(t) \in B(\mathcal{H})$ defined by $(S(t)h)(x) = h(x+t)$ is strongly continuous, and
4. the binary operation $(h \star g)(x) = h(x) \int_0^x g(u)du$ is well defined on (a subspace if necessary) $\mathcal{H}$.

Condition 1 enables us to actually calculate the forward rate, since $\mathcal{H}$ is not a space of equivalence classes.

**Definition 4.1.** *Let $w : [0,\infty) \to (0,\infty)$ be an increasing function such that $\int_0^\infty w^{-1}(x)dx < \infty$. Define the weighted Sobolev space with respect to $w$ as*

$$\mathcal{H}_w = \left\{ h : \mathbb{R}_+ \to \mathbb{R} \,|\, h \text{ is absolutely continuous, and } \int_0^\infty h'(u)^2 w(u)du < \infty \right\},$$

*where $h'$ stands for the weak derivative of $h$.*

**Lemma 4.2.** *Let $\mathcal{H}_w$ have inner product*

$$\langle g, h\rangle = g(0)h(0) + \int_0^\infty g'(u)h'(u)w(u)du.$$

*Then $\mathcal{H}_w$ is a Hilbert space which satisfies conditions 1 to 3.*

*Proof.* The form $\langle \cdot, \cdot \rangle$ is clearly bilinear. To see that it in fact is an inner product, assume that $\langle h, h\rangle = 0$. Then $h(0) = 0$ and $\int_0^\infty h'(x)^2 w(x)dx = 0$. Since $w(x)$ is positive, $h' = 0$ $dx$-a.s. so that $h$ is constant. But $h(0) = 0$, so $h = 0$.

To see that $\mathcal{H}_w$ is a Hilbert space, let $\{h_n\}$ be a Cauchy sequence in $\mathcal{H}_w$. As $w$ is increasing, define the Lebesgue-Stieltjes measure $\lambda_w(B) := \int_B w(u)du$ on the Borel sets $B$ of $\mathbb{R}_+$. Then the derivative sequence $\{h_n'(\cdot)\}$ can be embedded into a Cauchy sequence in $L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda_w)$, so that there exists a limit $h_0 \in L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda_w)$. Define

$$h(x) := \lim_{n\to\infty} h_n(0) + \int_0^x h_0(u)du.$$

Then $h$ is absolutely continuous and $h_n \to h$ in $\mathcal{H}_w$. It remains to check that $\mathcal{H}_w$ satisfies conditions 1 to 3.

1. For $h \in \mathcal{H}_w$, $h(x) = h(0) + \int_0^x h'(u)du$, so by Hölder's inequality,

$$|h(x)| \leq |h(0)| + \int_0^x |h'(u)|du = |h(0)| + \int_0^x |h'(u)|w(u)^{1/2}w(u)^{-1/2}du \leq$$
$$|h(0)| + \left(\int_0^x w(u)^{-1}du\right)^{1/2}\left(\int_0^x h'(u)^2 w(u)du\right)^{1/2} \leq$$
$$|h(0)| + \left(\int_0^\infty w(u)^{-1}du\right)^{1/2}\left(\int_0^\infty h'(u)^2 w(u)du\right)^{1/2}.$$

Then

$$h(x)^2 \leq 2\left(h(0)^2 + \left(\int_0^\infty w(u)^{-1}du\right)\left(\int_0^\infty h'(u)^2 w(u)du\right)\right) \leq$$
$$2\max\left\{1, \int_0^\infty w(u)^{-1}du\right\}\|h\|^2,$$

so $\delta_x$ is continuous. Notice here that the right-hand side is independent of $x$, so that in fact, by the Banach-Steinhaus theorem,

$$\sup_{x\in\mathbb{R}_+} \|\delta_x\| < \infty.$$

2. By the above discussion, it follows easily that

$$|I_x(h)| \le x \sup_{u\in\mathbb{R}_+} |\delta_u(h)| \le x\|h\| \sup_{u\in\mathbb{R}_+} \|\delta_u\|.$$

3. To see that $S(t)$ is strongly continuous, first notice that for every $h \in H_w$, the family $\{|S(t)h'|^2 \,|\, t \in \mathbb{R}_+\}$ is uniformly integrable with respect to $\lambda_w$: Let $\epsilon > 0$ and choose $K$ such that

$$\int_{\{|h'|^2>K\}} |h'|^2 d\lambda_w < \epsilon.$$

Then, substituting $u = x + t$ gives

$$\int_{\{x\in\mathbb{R}_+ \,|\, |h'(t+x)|^2>K\}} |h'(t+x)|^2 d\lambda_w(x) =$$

$$\int_{\{u\in[t,\infty) \,|\, |h'(u)|^2>K\}} |h'(u)|^2 w(u-t)du.$$

Since $w$ is increasing, this is dominated by

$$\int_{\{u\in\mathbb{R}_+ \,|\, |h'(u)|^2>K\}} |h'(u)|^2 w(u)du < \epsilon$$

proving that the family is uniformly integrable. As $h$ is absolutely continuous, it follows that $h'(x+t) \to h'(x)$, $\lambda$-a.s. and by the uniform integrability,

$$\int_0^\infty |h'(x+t) - h'(x)|^2 w(x)dx \to 0$$

as $t \to 0$. Then

$$\|S(t)h - h\|^2 = |h(t) - h(0)|^2 + \int_0^\infty |h'(x+t) - h'(x)|^2 w(x)dx \to 0$$

since $h$ is continuous.

□

Notice how condition number 2 follows directly from condition number 1, i.e. condition 1 guarantees that number 2 is also valid independent of the underlying Hilbert space.
To get condition 4, a subspace of $\mathcal{H}_w$ and an extra condition on $w$ is needed. The proof of the following can be found in [CT06].

**Proposition 4.3.** *Define* $\mathcal{H}_w^0 = \{h \in \mathcal{H}_w \,|\, h(\infty) = 0\}$ *and assume that* $\int_0^\infty \frac{x^2}{w(x)}dx < \infty$. *Then the binary operation in condition 4 is well defined on* $\mathcal{H}_w^0$ *and there exists a constant* $C$ *such that*

$$\|h \star g\| \le C\|h\|\|g\|. \tag{31}$$

It is now readily guaranteed that there exists a solution of the generalized HJM-model.

**Theorem 4.4.** *Assume that the map* $(x \mapsto \sigma^k(t,h,x))$ *belongs to* $\mathcal{H}_w^0$ *for every* $k, t$ *and* $h \in \mathcal{H}_w^0$. *Also assume that there exists a constant* $K$ *such that*

- $\sum_{k=1}^{\infty} \|\sigma^k(t,h)\|^2 \leq K\|h\|^2$ ,
- $\sum_{k=1}^{\infty} \|\sigma^k(t,h) - \sigma^k(t,g)\|^2 \leq K\|h-g\|^2$, *and*
- $\sum_{k=1}^{\infty} \|\sigma^k(t,h) \star \sigma^k(t,h) - \sigma^k(t,g) \star \sigma^k(t,g)\| \leq K\|h-g\|$.

*Then there exists a unique mild solution to the generalized HJM-model in (29) on every weighted Sobolev space* $\mathcal{H}_w$

*Proof.* Define

$$\alpha : [0,T] \times \mathcal{H}_w \to \mathcal{H}_w$$

by $\alpha(t,h) = \sum_{k=1}^{\infty} \sigma^k(t,h) \star \sigma^k(t,h)$ and

$$\sigma : [0,T] \times \mathcal{H}_w \to L_2(U, \mathcal{H}_w)$$

such that $\sigma(t,h)f_k = \sigma^k(t,h)$. Equation (29) is equivalent to (30). The proof now follows by Theorem 3.4. ∎

# References


[Bar95] Robert G. Bartle. *The Elements of Integration and Lebesgue Measure*. Wiley Classics Library, 1995.

[Bob05] A. Bobrowski. *Functional Analysis for Probability and Stochastic Processes*. Cambridge University Press, 2005.

[BØP05] T. Zhang B. Øksendal and F. Proske. Backward stochastic partial differential equations with jumps and application to optimal control of random jump fields. *Stochastics, 77, No. 5, 381 - 399*, 2005.

[CT06] R. Carmona and M. Tehranchi. *Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective*. Springer, 2006.

[DU77] J. Diestel and J. J. Uhl. *Vector Measures*. American Mathematical Society, Providence, RI, 1977.

[Eva98] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, RI, 1998.

[KS98] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, Providence, RI, second edition edition, 1998.

[MFT94] Yoichi Oshima Masatoshi Fukushima and Masayoshi Takeda. *Dirichlet Forms and Symmetric Markov Processes*. de Gruyter Studies in Mathematics 19. Walter de Gruyter, 1994.

[MY07] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and Their Applications*. Lecture Notes in Mathematics. Springer, 2007.

[NKQ97] S. Peng N. Karoui and M.C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance, Vol. 7, No. 1, 1-71*, 1997.

[Øks05] B. Øksendal. *Stochastic Differential Equations: An Introduction With Applications*. Springer, 2005.

[PR07] C. Prevot and M. Röckner. *A Consise Course on Stochastic Partial Differential Equations*. Springer, 2007.

[PZ92] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics, Volume 45. Cambridge University Press, 1992.